

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2003, 10:00:48 ; Search time 258 Seconds
(without alignments)
3063.765 Million cell updates/sec

Title: US-09-856-836-2
Sequence: 1 MAARKQIPUCGCHTRVYVNDX.....EEIASNSDSTISSTPEVKA 351

Scoring table:
BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ .p2n.model -DEV=xlh
-O=/cgn2_1/USPFO.spool/US0985636/runal.31012003.144236.22173/app.query.fasta.1.519
-DB=N-Geneseq.101002 -OEMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPC=-0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human0.cdi
-LIST=45 -DOCALLIGN=200 -THR=SCORE=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0985636.ecgn.1.1.0.etrnal.31012003.144236.22173 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N-Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1832	99.6	1630	21	AAA49249
2	1779.5	96.7	1820	20	AAV59096
3	1758.5	95.6	1847	19	AAV54162
4	1737	94.4	2812	23	ABV22737
5	1737	94.4	2812	23	ABV28565
6	1347.5	73.2	1430	21	AACT8126
7	1334.5	72.5	15320	22	AAK78660
8	1328.5	72.2	15332	22	AAK78662
9	1052.5	57.2	1331	23	ABL06979
10	876.5	47.6	4891	23	ABL06978
11	803.5	43.7	4179	23	ABL09294
12	789.5	42.9	1753	21	AAAC43745
13	781.5	42.5	1410	21	AAAC47397
14	771	41.9	1540	21	AAAC47955
15	770	41.8	1538	21	AAAC46123
16	759.5	41.3	1410	21	AAAC38458
17	615	33.4	522	24	ABO58251
18	517.5	28.1	313	23	AAS58064
19	467	26.5	311	24	AAK53945
20	454.5	24.7	2901	23	ABL16514
21	445	24.2	469	24	ABL093455
22	427.5	23.2	901	23	ABL16515
23	426.5	23.2	3165	23	ABL03412
24	425	23.1	686	20	AAZ24536
25	425	23.1	686	21	AAK65775
26	425	23.1	686	24	ABL48994
27	424.5	23.1	1072	21	AAAC3747
28	368.5	20.0	1246	21	AAAC3847
29	368.5	20.0	1253	21	AAAC3847
30	368.5	20.0	1256	21	AAAC3847
31	357.5	19.4	1276	21	AAAC3404
32	356.5	19.4	507	21	AAAC47884
33	353.5	19.2	987	21	AAAC43473
34	321	17.6	1166	23	ABL14459
35	321	17.4	223	15	AAO76451
36	301	16.4	3398	23	ABL14458
37	289	16.2	1710	22	AAK94176
38	279	15.2	1512	22	AAK93563
39	274	14.9	3450	22	AAA91119
40	274	14.9	3714	22	AAA91117
41	274	14.9	3747	22	AAA91115
42	274	14.9	3747	22	AAA91118
43	274	14.9	3747	22	AAA91120
44	274	14.9	3747	22	AAA91121
45	274	14.9	3747	22	AAA91122

ALIGNMENTS

RESULT 1
AAA49249 standard: cDNA: 1630 BP.

XX AAA49249;

XX 25-SEP-2000 (first entry)

DE Serine-threonine kinase receptor associated protein (STRAP) DNA sequence.

XX Serine-threonine kinase receptor associated protein; STRAP; mouse;

KW transforming growth factor beta; TGF-beta; proliferation; tumour growth;

KW cytosolic; anti-arteriosclerotic; anti-diabetic; nephrotic; cancer;

XX atherosclerosis; diabetes; ss.

XX Mus sp.

OS

XX

XX

FH

key

Location/Qualifiers

```

FT CDS 86.1141
FT /*tag= "STRAP"
FT /product= "Serine-threonine kinase receptor associated
FT /note= "protein"
FT /transl_except= (pos:143..145, aa:Xaa)
FT /note= "Xaa = unknown"
FT /transl_except= (pos:431..433, aa:Xaa)
FT /transl_except= (pos:776..778, aa:Xaa)
FT /transl_except= (pos:1004..1006, aa:Xaa)
FT /note= "Xaa = Leu, Ile"
XX PN WO200034310-A1.
XX PD 15-JUN-2000.
XX 10-DEC-1999; 99WO-US29267.
XX 10-DEC-1998; 98US-0111668.
XX (UYVA-) UNIV VANDERBILT.
XX PI Data PK, Moses HL;
XX DR WPI; 2000-442141/38.
XX P-PSDB; AAB01250.
XX PT Serine-threonine kinase receptors associated protein and the
XX PT polynucleotide encoding the receptor, useful for treating disease
XX PT associated with transforming growth factor beta activity such as cancer
XX PS Claim 18; Page 129-130; 134pp; English.
XX CC This sequence encodes the murine serine-threonine kinase receptor
XX CC associated protein (STRAP). The STRAP polypeptide is capable of
XX CC modulating transforming growth factor beta (TGF-beta) activity. TGF-beta
XX CC belongs to a family of multifunctional cell regulatory factors, which
XX CC bind to cell surface receptors. In certain cells TGF-beta promotes cell
XX CC proliferation, and in others it suppresses proliferation. Some tumour
XX CC cells use TGF-beta as an autocrine growth factor. STRAP enhances the
XX CC anti-proliferative activity of TGF-beta. The invention includes isolated
XX CC and purified anti-STRAP antibodies, and a hybridoma cell line producing
XX CC the monoclonal antibody. STRAP exhibits cytosolic, antiarteriosclerotic,
XX CC antidiabetic, and nephrotropic activity. The STRAP gene can be used to
XX CC detect STRAP encoding RNA. Anti-STRAP antibodies are used to detect STRAP
XX CC polypeptides in a sample. Compositions comprising STRAP polypeptides can
XX CC be used to treat disorders associated with TGF-beta, such as cancer,
XX CC atherosclerosis and diabetes.
XX SQ Sequence 1630 BP; 455 A; 346 C; 406 G; 419 T; 4 other;
XX
Alignment Scores:
Pred. No.: 1,23e-196 Length: 1630
Score: 1832.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.57% Indels: 0
DB: 21 Gaps: 0
US-09-856-836-2 (1-351) x AAAA9249 (1-1630)
QY 1 MetalMetalArg1nHrProLeuThrCysSerGlyHisThrArgProValAsp*** 20
DB 86 ATGGCATGATGAGGAGACGCCGCTACCTGCTGGGCGACACGGGCGCGGTGGGAGNTG 145
QY 21 AlaphSerGlyIleThrProTyrGlyTyrPheLeuIleSeraIaCysLysAspGlyLys 40
DB 146 GCCTTCAGGCGGACACGCGCTTACGCTACTTCTGATCAGCGCTGCAGAAAGATGGCAAG 205
QY 41 PrometLeuArgInGlyAspThrGlyAspThrIleGlyIleThrPheLeuGlyHisLysGly 60
DB 206 CCCATCTCCGCGGAGGAGATACAGGAGACTGGATTGGAAACATTGTTGGTCAATAAGGT 265

```

```

QY 61 AlavAlrPrlYAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
DB 266 GCGTTTGGGGTGCACATTAATTAAGAGATGCCACCAAGCTGGACAGCAGCTGCAGAC 325
QY 81 PheThrAlaLysValITrpAspAlaValSerGlyAspGluMetITrLeuAlaHisLys 100
DB 326 TTCACAGCCCAAGATGAGTGGGCTCTCAGAGATGAATTGATGACCCCTGCATTAAG 385
QY 101 HisLeValLysThrValAspPheThrGlnAspSerAsnTyrLeu***TrnGlyGln 120
DB 386 CACATTGTCAAGACTGTGGATTTCACACAGATGCAATTAACCTGMACTGGGGGACAG 445
QY 121 AspLysLeuArgIleTyrAspLeuAsnLysSerProGluAlaGluProLysGluIleSer 140
DB 446 GATTAAGTCTCCGCATATATGACTTGAACAAACCTGAACAGACACTTAAGAAATAGT 505
QY 141 GlyHisThrSerGlyIleLysLysAlaLeuITrPcysSerAspLysGlnIleLeuSer 160
DB 506 GGCACACTTCTGTATTAAAGGCTCTGTGGTGCAGTACGATAACAGATCCTTTCA 565
QY 161 AlaAspAspLysThrValArgLeuITrPaspHisAlaThrMetITrGluValLysSerLeu 180
DB 566 GCGGATGATAAACTGTTCGCTGGATCAGATCCCAATGACAGAAAGTAAATCTCTG 625
QY 181 AsnPheAsnMetSerValSerSerMetGluTyrIleProGluGluIleLeuValIle 200
DB 626 AATTTATATATGCTCTGTACAGCATGAGATATATCTCAGAGAGATTTGGTTAT 685
QY 201 ThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluITrPheLysSerPhe 220
DB 686 ACTTATGAGAGATCTATGTCTTTCATATGACAGTAACTGTGGAGCCAAATTAATCCTTT 745
QY 221 GluAlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAla 240
DB 746 GAAGCTCTCGACATCAATTTCTGCTTMTTCATCAAGAGAGAGTTCTGTGGCG 805
QY 241 GlyLysGluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluGluLeuGluSer 260
DB 806 GGTGGAAGAACATTAAACGTACCAACTATGATTAACAGTGGAGAGAGATGAGATCC 865
QY 261 TyrLysGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyLeuTyr 280
DB 866 TACAAAGTCACTTGTCTCCATTCACCTGTGTGATTCAGTCTGATGGGGAACCTAT 925
QY 281 AlSerGlySerGluAspGlyThrLeuArgLeuITrPheValValGlyLysThrTyr 300
DB 926 GCCAGCGTCTGAGATGGGACATTGAGATGTGGCAAACTGTGGTAAAGAACCTAT 985
QY 301 GlyLeuITrPcysVal***ProGluGluAspSerGlyGluLeuAlaLysProLysIle 320
DB 986 GGCCTGTGGAATGCGTGMTCCTGAGGAAGACACGGGGGAGACGGAAGCCAAAGATC 1045
QY 321 GlyPheProGluThrAlaGluGluGluLeuAlaGluIleAlaSerGluAsnSerAsp 340
DB 1046 GGAATTCACAAACAGCAGAGAGAAAGACTGTGCAGAAATGCTTCAGAAATTCAGAT 1105
QY 341 SerIleTyrSerSerIThrProGluValLysAla 351
DB 1106 TCCATCTATTCACTCACTCTGAAAGTTAAGGCC 1138

```

```

RESULT 2
AAVS9096
ID AAVS9096 standard; DNA; 1820 BP.
XX
AC AAVS9096;
XX
DT 20-JAN-1999 (first entry)
XX
DE TGFAS protein coding sequence.
XX
KW TGFAS; transforming growth factor-beta receptor associated protein;
signal transduction; sepsis; toxic shock; autoimmune thyroiditis; asthma;

```



```

FH Key Location/Qualifiers
FT CDS /tag= a
FT /product= "SH-SY5Y protein"
XX
XX WO9839437-A1.
XX
XX 11-SEP-1998.
XX
XX PD 05-MAR-1998; 98WO-JP00905.
XX
XX PF 05-MAR-1998; 98WO-JP00905.
XX
XX PR 05-MAR-1997; 97JP-0050302.
XX
XX PA (KYOM ) KYOMA HAKKO KOGYO KK.
XX
XX PI Sakaki Y;
XX
XX DR WPI: 1998-495844/42.
XX
XX DR P-PSDB: AAM54161.
XX
XX PT Novel apoptosis-related DNAs and proteins - for diagnosis,
XX preventing or treating diseases associated with apoptosis.
XX
XX PS Claim 1; Pages 40-43; 70pp; Japanese.
XX
XX CC This is the nucleotide sequence of a nucleic acid of the human cell
XX line SH-SY5Y, used in the method of the invention involving the use of
XX novel apoptosis-related DNAs and proteins. The inventions can be used
XX as diagnostic reagents for apoptosis e.g. (monoclonal) antibodies for
XX the protein, as a reagent in immunohistological staining, as apoptosis
XX inhibitors. It can also be used for treatment of apoptosis-related
XX diseases.
XX
XX SQ Sequence 1847 BP; 513 A; 389 C; 422 G; 523 T; 0 other;

Alignment Scores:
Pred. No.: 2,92e-188 Length: 1847
Score: 1758.50 Matches: 338
Percent Similarity: 96.88% Conservative: 4
Best Local Similarity: 95.75% Mismatches: 8
Query Match: 95.57% Indels: 3
DB: 19 Gaps: 1

US-09-856-836-2 (1-351) x AAM54162 (1-1847)
OY 1 MetAlaMetArgSlnThrProLeuThrCysSerGlyHisThrArgProValAlaASP** 20
DB 295 ATGGCAATGAGACAGACGCCGCTGCTGCTGCGCACACAGGACCGGTTGATTG 354
OY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 355 GCCTTCAGTGGCATCAGCGCTTATGGTATTTCTTAATCAGCCCTTGCAAAAGATGTA 414
OY 41 PrometLeuArgGln-GlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysG 60
DB 415 CCATCTCTACGGCAGGAGATACAGAGACTGATTTGGAACATTTTGGCTATAAAGG 474
OY 60 yAlaValITrPGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAs 80
DB 475 TGCTGTGGGGTGCAACACTGATTAAGATGCGACCAAGACGCTACAGCTGCA 534
OY 80 PheThrAlaLysValITrPAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisL 100
DB 535 TTTCACAGCCAAAGTGGGATGCTGCTCAGAGATGATGATGATGACCTGCTATA 594
OY 100 SHSIIeValLysThrValAlaSPheThrGlnAspSerAsnTyrLeu**ThrGlyGlyG 120
DB 595 ACACATTGTCMAACACTGTGATTTCACGCGAGTAGTAATTAATTTATTGTTAACCGGGACA 654
OY 120 nAspLysLeuLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluLys 140
DB 655 GATTAACACTGTACGATATATGACTTTGAACAACCTGGAAGCAGAACCTTAAGAAATTAG 714

```

```

OY 140 rGlyHisThrSerGlyIleLysLysAlaLeuTrpCysSerAspAspLysGlnIleLeuSe 160
DB 715 TGCTCATACTTTCGTATATAAAAAAGCTCTGCTGCTGACGAGGAGATTAACAGATTCTTC 774
OY 160 rAlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLe 180
DB 775 TGCTGATGACMAAACTGCTTGACCTTGATCATGCTACTATGACAGMAAGTGAATCTCT 834
OY 180 uAsnPheAsnMetSerValSerSerMetGluTyrIleProGluGluValIleLeuVal 200
DB 835 AAATTTAAATATGCTGTTAGTATGATGAAATATATTCCTGAGGAGACATTTTGGTTAT 864
OY 200 eThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPh 220
DB 895 AACTTATGACGATTCATCTCTTTCATGATGACAGTATGATGACCAATTAATCTCT 954
OY 220 eGluAlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuVal 240
DB 955 TGAAGCTCCGCAACCATCATCTGCTCATCTCTCATCTGAGAGAAATTTCTTGTTC 1014
OY 240 aGlyGlyGluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluGluLeuGluSe 260
DB 1015 AGCGCGTGAAGATTTTAACTTTATAGTATGATTAATATAGTGAGACAAATTAGATC 1074
OY 260 rTyrLysGlyHisPheGlyProIleHisCysValArgPheSer- ProAspGlyGluLeu 280
DB 1075 CTCAACAGGACACTTTCCTCTCTATTCACGTGAGAGATTTAGTCCCTGATGAGAACTCT 1134
OY 280 yAlaAspSerGlySerGluAspGlyThrLeuArgLeuTrpGlnThrValGlyLysThr 300
DB 1135 ATCCCACTGGTTGAGAAAGATGAAACATTCATGACATTCGCAACTGTGTGAGAAAACGT 1194
OY 300 yGlyLeuTrpLysCysVal**ProGluGluAspSerGlyGluLeuAlaLysProLys 320
DB 1195 ATGGCCTTGGAAATATGCTGCTCTCTCAACAGATATGCTGACCTGCAACGCAAGA 1254
OY 320 leGlyPheProGluThrAlaGluGluGluLeuAlaGluIleAlaSerGluAsnSer 340
DB 1255 TTGGTTTTCAGAGACAAACAGACAGACTA--GAAGAAATGCTTCAGAGAAATTCAG 1311
OY 340 sPseIleTyrSerSerThrProGluValLysAla 351
DB 1312 ATTGCATCTTCTCTGCTGCTGATGTATAGGCC 1346

RESULT 4
ABV22737
ID ABV22737 standard; cDNA; 2812 BP.
XX
XX AC ABV22737;
XX
XX DT 13-SEP-2002 (first entry)
XX
XX DE Human prostate expression marker cDNA 22728.
XX
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN W0200160860-A2.
XX
XX PD 23-AUG-2001.
XX
XX PF 20-FEB-2001; 2001WO-US05171.
XX
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211310P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

```


prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

Claim 1; Page 5974-5975; 11750P; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV6213) of the specification or its complement. (I) is useful for:

(a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

(e) selecting a composition for inhibiting prostate cancer in a patient;

(f) assessing the prostate cell carcinogenic potential of a compound;

(g) determining whether prostate cancer has metastasized in a patient;

(h) assessing the aggressiveness or indolence of prostate cancer in a patient;

(i) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 2812 BP: 778 A; 614 C; 601 G; 819 T; 0 other:

Alignment Scores:

Pred. No.:	1.43e-185	Length:	2812
Score:	1737.00	Matches:	336
Percent Similarity:	96.33%	Conservative:	5
Best Local Similarity:	94.92%	Mismatches:	8
Query Match:	94.40%	Indels:	5
DB:	23	Gaps:	1

US-09-856-836-2 (1-351) x ABV28565 (1-2812)

QY 1 MetAlaMetArgGlnThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
 DB 336 ATGGCAATGACAGACAGCGCTCAGCTGCGCACACGACCGCCCTGGTGTGATTTG 395
 QY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
 DB 396 GCGTTCAGGCGATCAGCGCTATGCGTATTTCTTAATCAGCGCTGCAAAAGTGTAA 455
 QY 41 PrometLeuArgGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
 DB 456 CCTATGCTACGCGCAGGAGATACAGAGAGCTGATGGAACATTTTGGCTCAATAAGGT 515
 QY 61 AlaValThrGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrLysAlaAlaAsp 80
 DB 516 GGTGTTGGGGTGCACACTGATAGATGATGCCACCAAGCAGCTACAGCAGCTGACAT 575
 QY 81 PheThrAlaLysValTyrAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
 DB 576 TTCACAGCCAAAGTGGGATGCTGCTCAGGAGATGATGATGATGATGATGATGATG 635
 QY 101 HisIleValLysThrValAspPheThrGlnAspSerAspTyrLeu***ThrGlyGlyGln 120
 DB 636 CACATTTGTAACACTGTGATTTACCGCAGAGATGATGATGATGATGATGATGATG 695
 QY 121 AspLysLeuLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluLysSer 140
 DB 696 GATTAACCTTTCGATATATGATGATGATGATGATGATGATGATGATGATGATGATG 755
 QY 141 GlyHisThrSer-GlyTyrIleLysLysAlaLeuTyrCysSer-AspAspLysGln-IleLeu 159
 DB 756 GGTCACTACTTGGGTATTAATAAAAGCTGCTGTGTGCTGAGTGAACACAAATTCCT 815
 QY 160 SerAlaAspAspLysThrValArgLeuTyrPhePheHisAlaThrMetThrGlyValLysSer 179
 DB 816 TCTGCTGATGACAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 875
 QY 180 LeuAsnPheAsnMetSerValSerSerMetGluTyrIleProGluGlyGluIleLeuVal 199
 DB 876 CTAAATTTTAATATGCTGTGTAGTACTATGATGATGATATATTCCTGAGGAGAGATTTGGTT 955

QY 200 IleThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSer 219
 DB 936 ATAACTATAGGACATATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 995
 QY 220 PheGluAlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuVal 239
 DB 996 TTGGAAGCTCCGCAACCATCAATCTGCATCTTCATCCCGAGAAAGAAATTTCTTGT 1055
 QY 240 AlaGlyGlyGluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluLeuGlu 259
 DB 1056 GAGGCGGGGAAAGATTTTAACCTTATATGATATATATATATATATATATATATATAT 1115
 QY 260 SerTyrLysGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyLeu 279
 DB 1116 TCTTCAAGGAGACATTTGCTCTTATCTGATGATGATGATGATGATGATGATGATG 1175
 QY 280 TyrAlaSerGlySerGluAspGlyThrLeuAlaGlyLeuThrGlnThrValAlaGlyLysThr 299
 DB 1176 TATGCCAGCTTCAGAAACATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1235
 QY 300 TyrGlyLeuTrpLysCysVal***ProGluGluAspSerGlyGluLeuAlaLysProLys 319
 DB 1236 TATGCGCTTTCGAAATGATGTGCTTCTGAAAGAAATGATGATGATGATGATGATG 1295
 QY 320 IleGlyPheProGluThrAlaGluGluGluLeuAlaGluGluIleAlaSerGluAsnSer 339
 DB 1296 ATTTGTTTCCAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1349
 QY 340 AspSerIleTyrSerSerThrProGluValLysAla 351
 DB 1350 GATTCATCT 1385
 RESULT 6
 AAC78126
 ID AAC78126 standard; cDNA; 1430 BP.
 AC AAC78126;
 XX 08-FEB-2001 (first entry)
 XX Human cancer associated gene sequence SEQ ID NO:520.
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiaslmetlic; antirheumatic; antihistilic; antiviral;
 KW antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; noctropic;
 KW vasotropic; antipruritic; antianogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX Homo sapiens.
 OS WO20005350-A1.
 PN 21-SEP-2000.
 PD 08-MAR-2000; 2000MO-US05882.
 PF 12-MAR-1999; 99US-0124270.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM;
 PI MPI, 2000-587533/55.
 DR P-PDB; AAB3917.
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -

PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0239678.
 XX
 XX (HUMA-) HUMA GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 XX metastasis -
 XX
 PS Disclosure; SEQ ID NO 33472; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 15320 BP; 4488 A; 2687 C; 3177 G; 4968 T; 0 other;

Alignment Scores:

Pred. No.:	4,43e-139	Length:	15320
Score:	1334.50	Matches:	291
Percent Similarity:	66.81%	Conservative:	13
Best Local Similarity:	63.96%	Mismatches:	44
Query Match:	72.53%	Indels:	110
DB:	22	Gaps:	3

US-09-856-836-2 (1-351) x AAK78660 (1-15320)

PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0228343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251989.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 PT WPI, 2001-483426/52.
 PT XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PT XX
 PS Disclosure: SEQ ID NO 33474; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 CC XX
 SQ Sequence 15332 BP; 4490 A; 2686 C; 3180 G; 4976 T; 0 other;

Alignment Scores:

Pred. No.:	2,11e-138	Length:	15332
Score:	1328.50	Matches:	290
Percent Similarity:	66.59%	Conservative:	13
Best Local Similarity:	63.74%	Mismatches:	45
Query Match:	72.20%	Indels:	110
DB:	22	Gaps:	3

US-09-856-836-2 (1-351) x AAK78662 (1-15332)

Percent Similarity: 75.79%
 Best Local Similarity: 60.06%
 Query Match: 57.20%
 DB: 23 Gaps: 1

US-09-856-836-2 (1-351) x ABL06978 (1-1331)

```

OY 3 MetArgInThrProLeuThrcCysSerGlyHisThrArgProValValAsp***AlaPhe 22
DB 360 TTGGCAACAAATCCCTGACCTGACAGTGGCCACAGCAGCCGCTGCTCCACTGACCTTC 419
OY 23 SerGlyIleThrProGlyIleThrPheLeuIleSerAlaCysLysAspGlyLysProMet 42
DB 420 AGCGACATCTGCGACGCGGCTACTTCTCATCTGCGCTGCAAGATGGCAGCCGAG 479
OY 43 LeuArgGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGlyAlaVal 62
DB 480 CTGCGGACGCGGACAGCGGCTGCGTGGCCACTTTCGACCTTCGACACACAAAGGCGCTG 539
OY 63 TrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThr 82
DB 540 TGGAAAGCGCAGCCTGACCGAAATGCAACCCCTGCGGCGGCGGCGGCGGCGGCGGCGG 599
OY 83 AlaLysValTrpAspAlaValSerGlyLysGlyLeuMetThrLeuAlaHisLysHisIle 102
DB 600 GCGAAGTGTGGAAATGCGGCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 659
OY 103 ValLysThrValAspPheThrGlnAspSerAsnTrpLeu***ThrGlyGlyGlnAspLys 122
DB 660 GTGAAGAGGCTGCGCTTCACAGAGGATTCGAGAACATGTCGTCGCGCAGCAACGAGAG 719
OY 123 LeuLeuArgIleTrpAspLeuAsnLysProGlnAlaGluProLysGluIleSerGlyHis 142
DB 720 CTGGTGGCGGCTTCAACCTGAGCAGCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 779
OY 143 ThrSerGlyIleLysLysAlaLeuTrpCysSerAspAspLysGlnIleLeuSer---Ala 161
DB 780 ACGGGGCGCATCAAGCGGCGCTCTTCTCCCGCGGACAAAGCAGTCATCTCGCGCGCC 839
OY 162 AspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGlnValLysSerLeuAsn 181
DB 840 GAGGACAAACAGCGTGGCGCTTCGCGACCCGATGACGCGGCGGATGAGGTCGACCGCC 899
OY 182 PheAsnMetSerValSerSerMetGluTrpIleProGluGlyGluLeuValIleThr 201
DB 900 TTCACAGCAATCCGACAGCTAGAGATCTCCAGGACACACATCTGACCATATGCG 959
OY 202 TrpGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProLeyLysSerPheGlu 221
DB 960 CACGGCTCTCGATTAAGCTTCTGCGAGATCGACACGCTGAAGAGCTGAAGGAGTGAAA 1019
OY 222 AlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAlaGly 241
DB 1020 GTGCCACGACAGCTGGCGCTGCGGCGCTTTCGATCCGCAATGAGATGTTGTTGTGGC 1079
OY 242 GlyLysAspPheLysLeuTrpLysTrpAspTrpAsnSerGlyGluLeuGluSerTrp 261
DB 1080 GCGGAGAGCTTAAGATGATGACAAATTCGACTACATCACAGAAACGAAATTTGATCTTC 1139
OY 262 LysGlyHisPheGlyProLeyHisCysValArgPheSerProAspGlyGluLeuTrpAla 281
DB 1140 AAGGCAATTTGGCGCCCTGCGACAGCTGAAATTCAGCCCGGATCGGCAACTCTATGCC 1199
OY 282 SerGlySerGlyAspGlyThrLeuArgLeuTrpGlnThrValValAlaGlyLysThrTrpGly 301
DB 1200 ACGGCTCGAGGAGCGGCGCTCTCGGCTGTGGCAGACACACCGTGGGCAAGACTTACGCC 1259
OY 302 LeuTrpLysCysVal***ProGluLysAspSerGlyGluLeuAlaLysProLys 319
DB 1260 TTGTGAAATGACAGGAGCCGCGGACCTGGCAATTCATCATCTCGCGCGC 1313
  
```

```

ID ABL06978 standard; cDNA; 4891 BP.
XX
AC ABL06978;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 15416.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
PI MPI; 2001-656860/75.
XX
PS P-PSDB; ABB62875.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
Claim 1: SEQ ID NO 15416; 21pp + Sequence Listing; English.
XX
XX
The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
capable of detecting 1000 or more genes from Drosophila. The invention is
cell-cell interactions in higher eukaryotes for the development of
cell-cell interactions, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
sequences (AB16176-AB16175) and the encoded proteins
(AB16176-AB16175).
XX
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp://wipo.int/pub/publ/publseq.pct_sequences.
XX
XX
SQ Sequence 4891 BP; 1166 A; 1296 C; 1227 G; 1202 T; 0 other;
XX
XX
Alignment Scores:
Pred. No.: 4.31e-88 Length: 4891
Score: 876.50 Matches: 176
Percent Similarity: 63.99% Conservative: 55
Best Local Similarity: 48.75% Mismatches: 77
Query Match: 47.64% Indels: 55
DB: 23 Gaps: 4
  
```

```

US-09-856-836-2 (1-351) x ABL06978 (1-4891)
OY 38 AspGlyLysProMetLeuArgGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGly 57
DB 1929 GATGGCAACCCCGATGCTGCGGCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1870
OY 58 HisLysGlyAlaValAlaTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAla 77
DB 1869 CACAAGGCGCGCGTGTGGAAACGCCACGCTGAACCGAAATGCGACCCCTGCGGCGG 1810
OY 78 AlaAlaAspPheThrAlaLysValTrpAspAlaValSerGlyLysGlyLeuMetThrLeu 97
DB 1809 GCGGCGGCTTCAACGCGGCAAGGTGTGGAAATGCGGCGGCGGCGGCGGCGGCGG 1750
OY 98 AlaHisLysHisIleValLysThrValAspPheThrGlnAspSerAsnTrpLeu***Thr 117
  
```


QY 171 HisAlaThrMetThrGluValIysSerIleuAsnPheAsnMetSerValSerSerMetGlu 190
||| |||||: ||| ||||| :||| :|||: |||||
Db 3818 CGCATGACGGGATGCGAGCCGCTGACGATTCACGACATCCGACACCTAGAG 3759
QY 191 TyrIleProGluGluIleuValIleuThrGlyArgSerIleuAlaPheHisSer 210
||| ||||| :||| ||||| :||| ||||| :|||
Db 3758 ATCTCCAGGACACACACATGACGACATGCGACGCGCTCCGATTTAGCTTCGGAG 3699
QY 211 AlaValSerIleuGluProIleuLysSerPheGluAlaProAlaThrIleuAsnSer 230
: ||| :||| :||| :||| :||| :||| :||| :|||
Db 3698 ATCGACACGCTGACGAGCGTGAAGGTGAAGTCCGACGACGCTGGCGCAGT 3639
QY 231 ***HisProGluIuSGluPheLeuValAlaGlyGluAspPheLysLeuTyrLysTyr 250
||||| :||| :||| ||||| ||||| :||| :||| :||| :||| :||| :||| :|||
Db 3638 TTGCATCCGGATAGCATGTGTTGTTGTGGCGGCGAGGACTTTAAGATGACAAATTC 3579
QY 251 AspTyrAsnSerGlyGluIleu-----GluSerTyrLysGlyHisPheGlyProIleHis 269
||||| :||| :||| ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3578 GACTACATCACAGGAACGAAT-TGCTATGACTTTCACATGCGTCTTAATTC 3520
QY 259 -----GluSerTyrLysGlyHisPheGlyProIleHis 269
||||| :||| :||| ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3519 CAAGGACGTGTCTTCTGTTCCGCCACAGAAATCTTCAGGACATTCGGCGCCGTCAC 3460
QY 270 CysValArgPheSerProAspGlyGluLeuTyrAlaSerGlyAspGlyThrIleu 289
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3459 AGCGTAAAGTTGACCCCGGATGCGCACTTANCCAGCGGCTCCGAGAGCGGACTCTC 3400
QY 290 ArgIleuTrpGlnThrValAlaGlyLysThrTyrGlyLeuTrpLysCysVal***ProGlu 309
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3399 CGGCTGTGGAGACACCGGCGGCAAGACTACGGCTTGGAAAGTCACGAGCCCGCG 3340
QY 310 GluAspSerGlyIleuAlaLysProLys----- 319
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3339 GACCTGGGCAATTCATCAGCTGCGCGGAGGTGCAAGCAAACTGATGTGGCGAGC 3280
QY 320 -----IleGlyPhe 322
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3279 AGGAGTCGCTGTACACAGCCGACGCGGCGACGACGAAAGGACATGTATTC 3220
QY 323 ProGluThrAlaGluIleuAlaGlu-----IleAlaSerGluAsn 338
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3219 GGAGTCAGCGCTGATCCGATTCAGACAGAACAACTCGTAGTCGGCTAC-AAAT 3161
QY 339 SerAspSerIleTyrSerSer 345
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3160 AACGATCTTCATATTCACAC 3140
RESULT 12
AAC43745
ID AAC43745 standard; DNA; 1753 BP.
AC AAC43745;
XX
XX 18-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 40353.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
XX Zea mays subsp. mays.
OS
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

```

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145092.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148368.
PR 17-AUG-1999; 99US-0148175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154038.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

```

```

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:

```

Pred. No.: 6,54e-79
Score: 789.50
Length: 1753
Percent Similarity: 68.62%
Matches: 147
Best Local Similarity: 50.69%
Conservative: 52
Query Match: 42.91%
Mismatches: 88
DB: 21
Indels: 3
Gaps: 2

```

US-09-856-836-2 (1-351) x AAC43745 (1-1753)

```

QY 7 ProleuthrCysSerGlyHisThrArgProValAlaAsp***AlaPheSerGlyIleThr 26
DB 347 CCGCTGTGCTGCGCAGCGCCACTCGCGCGCGGTGGACCTCTTCTACAGCCCGCTACA 406
QY 27 ProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLysPrometLeuArgGly 46
DB 407 CCGGATGGCTACTCTTCATCATCAGCGCCAGCAAGGATACAAATCCAAATGCTGTAATG 466
QY 47 AspThrGlyAspTyrPheGlyThrPheLeuGlyHisLysGlyAlaValTyrGlyAlaThr 66
DB 467 GAGACTGGCGACTGGATGGAGCTTTTCAAGTCATAAAGAGCTGTTGGAGCTGTGC 526
QY 67 LeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAlaAspPheThrAlaLysValTyr 86
DB 527 CTTGACAGCAATGCTGTGAGGCGCATCTCTCTCTGCTGCTTACCTTAAGTAATGG 586
QY 87 AspaIaValSerGlyAspGlyLeuMetThrLeuAlaHisLysHisIleValLysThrVal 106
DB 587 GATGCACTTACTGGGAGGAGCTACTCATTTGATGACAAAGCATATAGTGGCGGTG 646
QY 107 AspheThrGlnAspSerAsnTyrLeu***ThrGlyGlyGlnAspLysLeuLeuArgIle 126
DB 647 GCCTTCTCTGAGATACACACACTTTTACTCACTGAGAGCATGGAAGATTTGGCTGTG 706
QY 127 TyrAspLeuAsnLysProGlnAlaGluProLysGlnLysSerGlyHisThrSerGlyIle 146
DB 707 TATGATATGAAATCGTCTGTAAGCTCAAGAGAGCTTACAAATACCTGTTCCGTT 766
QY 147 LysLysAlaLeuTyrCysSerAspAspLysGlnIleLeuSerAla-----AspaLys 164
DB 147 LysLysAlaLeuTyrCysSerAspAspLysGlnIleLeuSerAla-----AspaLys 164

```

```
Db 767 CGAAGCTGCTGCTGCTTCTAGTACCAACTATATAGTTCCTGCTGCTGCTGCTGCT 826
Oy 165 ThValArgLeuTrpAspHisAlaTrhMetThrGluValLysSerLeuAsnPheAsnMet 184
    |||||
Db 827 GGAGTAAAGGCTGGGATGTGAGAACTGCAAAAAAATGTCACAACTCTGAAACCAAGCT 886
Oy 185 SerValSerSerMetGluTrpLeuProGluGlyGluLeuValIleThrTrpGlyArg 204
    |||||
Db 887 TCTGTCTACCTGATGACAGCAAGCCAGATGCGATGCTACTGCTGCTGCTGCTGCTGCTGCT 946
Oy 205 SerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPheGluAlaProAla 224
    |||||
Db 947 AGCTATAAATCTGGGAGGCAATCACTTGGCTTGTAAACCTATGATATCCATGCTC 1006
Oy 225 ThrIleAsnSerAlaSer***HisProGluLys---GluPheLeuValAlaGlyGlyL 243
    ::|||
Db 1007 AAGTGTGAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066
Oy 244 AspPheLysLeuTrpLysTrpAspTrpAsnSerGlyLysGluLeuGluLysSerTrpLysGly 263
    |||||
Db 1067 GATTATGCGTTCATGATTTGATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
Oy 264 HisPheGlyProIleHisCysValArgPheSerProAspGlyGluLeuTrpAlaSerGly 283
    |||||
Db 1127 CATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1186
Oy 284 SerGluAspGlyThrLeuArgLeuTrpGln 293
    |||||
Db 1187 TCAGAGATGCGCACCATTCGATCTGCGAG 1216

RESULT 13
AAC47397
ID AAC47397 standard; DNA: 1410 BP.
XX AAC47397;
AC
XX 18-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 53661.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
XX EPI033405-A2.
PN
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
```

PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145222.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145919.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150864.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151348.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 13-SEP-1999; 99US-0153070.
 PR 15-SEP-1999; 99US-0153758.
 PR 16-SEP-1999; 99US-0154018.
 PR 20-SEP-1999; 99US-0154039.
 PR 22-SEP-1999; 99US-0154779.
 PR 23-SEP-1999; 99US-0155139.
 PR 24-SEP-1999; 99US-0155486.
 PR 28-SEP-1999; 99US-0155659.
 PR 29-SEP-1999; 99US-0156458.
 PR 04-OCT-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 21-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
 Pred. No.: 3,82e-78
 Score: 781.50
 Length: 1410
 Percent Similarity: 66.028
 Matches: 148
 Best Local Similarity: 47.90%
 Conservative: 56
 Query Match: 42.47%
 Mismatches: 94
 Indels: 11
 Gaps: 4

US-09-856-836-2 (1-351) x AAC47397 (1-1410)

QY 7 ProLeuThrcysSerGlyHisThrArgProValAlaAsp**AlaPheSerGlyIleThr 26
 DB 258 CCGTTAGTGTGCGACGGGACATTCGTCGCCGTGATGCTTATATATCGGTGACT 317

QY 27 ProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyIleThrPheMetLeuArgGly 46
 DB 318 CCGATGGGTCTTCTCATTTAGCGCCAGCAAGATTCGAATCGATGGTGAAGAAATGGA 377

QY 47 AspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGlyAlaValTrpGlyAlaThr 66
 DB 378 GAGACGTGTGACGGATGGGACATTTGAGGACATTAAGGACAGTTCGATTCGACAGC 437

QY 67 LeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTrp 86
 DB 438 CTTCGATTAATAATGCTATTCCTGCTGCCACGCTTCGATTCGATTCGATTCGATTCG 497

QY 87 AspAlaValSerGlyAspGlyLeuMetThrLeuAlaHisLysHisIleValLysThrVal 106
 DB 498 AATGCAATTCGACGAGATGATTCACCTCTTGAACACAGCAATTCGTCGATTCGATTCG 557

QY 107 AspPheThrGlnAspSerAsnTyrLeu**ThrGlyGlyClnAspLysLeuArgIle 126
 DB 558 GCCTTTCTGAGGACACTCCGTTTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617

QY 127 TyrAspLeuAsnLysProGluAlaGluProLysGluIleSerGlyHisThrSerGlyIle 146
 DB 618 TTGCAATTTGAATCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 677

QY 147 LysLysAlaLeuTrpCysSerAspAspLysGlnIleLeuSerAla-----AspAspLys 164
 DB 678 AGAAGCTGCAATGCTTCTTACTAGTATTAATACATTTAACTGCTTGAACAGATTCGCT 737

QY 165 ThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeuAsnPheAsnMet 184
 DB 738 GACATTAAGTTATGGGACAAAGACATGATTCCTTAACTGCTTGAACAGATTCGCTTGA 797

QY 185 SerValSerSerMetGluTyrIleProGluGlyGluIleLeuValIleThrTyrGlyArg 204
 DB 798 CCAATTACTAGTCTGCAAGTAACTGCAAGTAACTGCAAGTAACTGCAAGTAACTGCAAG 857

QY 205 SerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPheGluAlaProAla 224
 DB 858 AGTGTTAAGTTTGGGACGCTAAATTTTGGATTCCTGTAAGTAACTGCTGCTGCTGCTG 917

QY 225 ThrIleAsnSerAlaSer**HisProGluLys--GluPheLeuValAlaGlyGlyGlu 243
 DB 918 AATGTTGAATCGGACATCGTTGGAACGGAACGGAACGGAACGGAACGGAACGGAACGGA 977

QY 244 AspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluGluLeuGluSerTyrLysGly 263

Db	978	GATGTGGGGTGCACAGGTTGATTTCACAGCTCGGAGGAAAGATTTGGGTACCAAGGCT	1037
QY	264	HisPheGlyProIleHisCysValArgPheSerProAspCylGluLeuTYRAlaSerGly	283
Db	1038	CATCAGGGCCAGTGCACCTCGTGAGGTATGCCCCAGGAGGGAGTCATACCTCAGGC	1097
QY	284	SerLysAspGlyThrLeuArgLeuTyrPGLInThValValGlyLysThrTyrolLeuTyr	303
Db	1098	TCAGAGACGGAGGCGACAAATATGC-----GTGGTGGCTTCGGTAAAC-----	1142
QY	304	LysCysVal**ProGluLysAspSer	312
Db	1143	-----CATCCTGAGAGACCAAT	1160
RESULT 14			
MAC47955			
ID	MAC47955	standard; DNA: 1540 BP.	.
XX	AC		
XX	MAC47955;		
XX	DE	18-OCT-2000	(first entry)
XX	DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 55727.	
XX	KW	Hybridisation assay; genetic mapping; gene expression control;	
XX	KW	protein identification; signal transduction pathway;	
XX	KW	metabolic pathway; promoter; termination sequence; ss.	
XX	OS	Arabidopsis thaliana.	
XX	PN	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	PF	25-FEB-2000; 2000EP-0301439.	
XX	PR	25-FEB-1999;	9905-0121825.
PR	05-MAR-1999;	9905-0123180.	
PR	09-MAR-1999;	9905-0123548.	
PR	23-MAR-1999;	9905-0125788.	
PR	25-MAR-1999;	9905-0126264.	
PR	29-MAR-1999;	9905-0126785.	
PR	01-APR-1999;	9905-0127462.	
PR	06-APR-1999;	9905-0128234.	
PR	08-APR-1999;	9905-0128714.	
PR	16-APR-1999;	9905-0129845.	
PR	19-APR-1999;	9905-0130077.	
PR	21-APR-1999;	9905-0130449.	
PR	23-APR-1999;	9905-0130510.	
PR	23-APR-1999;	9905-0130891.	
PR	28-APR-1999;	9905-0131449.	
PR	30-APR-1999;	9905-0132407.	
PR	30-APR-1999;	9905-0132486.	
PR	04-MAY-1999;	9905-0132485.	
PR	05-MAY-1999;	9905-0132486.	
PR	06-MAY-1999;	9905-0132487.	
PR	07-MAY-1999;	9905-0132863.	
PR	11-MAY-1999;	9905-0134256.	
PR	14-MAY-1999;	9905-0134218.	
PR	14-MAY-1999;	9905-0134221.	
PR	14-MAY-1999;	9905-0134370.	
PR	18-MAY-1999;	9905-0134768.	
PR	19-MAY-1999;	9905-0134941.	
PR	20-MAY-1999;	9905-0135124.	
PR	21-MAY-1999;	9905-0135353.	
PR	24-MAY-1999;	9905-0136021.	
PR	27-MAY-1999;	9905-0136382.	
PR	28-MAY-1999;	9905-0136782.	
PR	01-JUN-1999;	9905-0137222.	

PR	05-AUG-1999	99US-01473203
PR	06-AUG-1999	99US-01474161
PR	09-AUG-1999	99US-01474933
PR	10-AUG-1999	99US-01479335
PR	11-AUG-1999	99US-01481719
PR	12-AUG-1999	99US-01483451
PR	13-AUG-1999	99US-01485651
PR	14-AUG-1999	99US-01486644
PR	16-AUG-1999	99US-01493638
PR	17-AUG-1999	99US-01491757
PR	18-AUG-1999	99US-01494265
PR	20-AUG-1999	99US-01497222
PR	20-AUG-1999	99US-01497233
PR	20-AUG-1999	99US-01499299
PR	23-AUG-1999	99US-01499299
PR	23-AUG-1999	99US-01499300
PR	25-AUG-1999	99US-01499300
PR	26-AUG-1999	99US-01508844
PR	27-AUG-1999	99US-01510065
PR	27-AUG-1999	99US-01510666
PR	30-AUG-1999	99US-01510800
PR	31-AUG-1999	99US-01513033
PR	01-SEP-1999	99US-01514348
PR	01-SEP-1999	99US-01515930
PR	07-SEP-1999	99US-01515930
PR	10-SEP-1999	99US-01532633
PR	13-SEP-1999	99US-01537080
PR	15-SEP-1999	99US-01540318
PR	16-SEP-1999	99US-01541019
PR	20-SEP-1999	99US-01547779
PR	22-SEP-1999	99US-01551319
PR	23-SEP-1999	99US-01554546
PR	24-SEP-1999	99US-01555689
PR	28-SEP-1999	99US-01564548
PR	29-SEP-1999	99US-01565116
PR	04-OCT-1999	99US-01571517
PR	05-OCT-1999	99US-01575733
PR	06-OCT-1999	99US-01576655
PR	07-OCT-1999	99US-01580229
PR	08-OCT-1999	99US-01582022
PR	12-OCT-1999	99US-01583639
PR	13-OCT-1999	99US-01592525
PR	13-OCT-1999	99US-01592523
PR	14-OCT-1999	99US-01593311
PR	14-OCT-1999	99US-01593637
PR	14-OCT-1999	99US-01595648
PR	18-OCT-1999	99US-01595654
PR	21-OCT-1999	99US-01607141
PR	21-OCT-1999	99US-01607657
PR	21-OCT-1999	99US-01607658
PR	21-OCT-1999	99US-01607760
PR	21-OCT-1999	99US-01608114
PR	21-OCT-1999	99US-01608315
PR	22-OCT-1999	99US-01609800
PR	22-OCT-1999	99US-01609881
PR	22-OCT-1999	99US-01609881
PR	25-OCT-1999	99US-01614004
PR	25-OCT-1999	99US-01614005
PR	26-OCT-1999	99US-01614006
PR	26-OCT-1999	99US-01613559
PR	26-OCT-1999	99US-01613659
PR	28-OCT-1999	99US-01613651
PR	28-OCT-1999	99US-01613922
PR	28-OCT-1999	99US-01614943
PR	29-OCT-1999	99US-01621122

OY 323 Proglutrhalauglu 328
||| |||
Db 1307 GATGAGCTTCCTCCTAG 1324
RESULT 15
AAC46123
ID AAC46123 standard; DNA; 1538 BP.
XX AAC46123;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48986.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
XX
PF 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123588.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144336.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2003, 10:00:47 ; Search time 71 Seconds

(without alignments)
7040.607 Million cell updates/sec

Title: US-09-856-836-1

Perfect score: 1630

Sequence: 1 ttaccgcctacccgtaggga.....atgcctctaataaataaa 1630

Scoring table: IDENTITY_NUC

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5H_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5H_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCUTS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1011.4	62.0	1820	2	US-08-828-922-2
2	240.8	14.8	686	4	US-09-040-984-36
3	240.8	14.8	686	4	US-09-123-912-36
4	240.8	14.8	686	4	US-09-643-597-36
5	50.6	3.1	7218	1	US-08-232-463-14
6	48.2	3.0	577	4	US-09-385-982-449
7	48.2	3.0	647	4	US-08-385-982-323
8	45.8	2.8	2152	1	US-08-188-582-17
9	45.8	2.8	2152	1	US-08-646-715-17
10	44.2	2.7	1221	3	US-08-965-600-2
11	44.2	2.7	1221	3	US-09-489-506-2
12	39.6	2.4	2085	2	US-08-283-917-8
13	39.6	2.4	2085	2	US-08-961-716-8
14	37.8	2.3	1607	2	US-08-883-534-2
15	37.8	2.3	1607	3	US-09-204-764-2
16	37.6	2.3	7042	4	US-09-092-508-1
17	37.6	2.3	7042	4	US-09-435-115-1
18	37.6	2.3	7042	4	US-09-098-310-1
19	37.6	2.3	7042	4	US-09-690-364-21
20	37.6	2.3	7042	4	US-09-093-508-15
21	37.6	2.3	7042	4	US-09-435-115-15
22	37.6	2.3	7042	4	US-08-936-165A-63
23	36.8	2.3	7811	2	US-08-751-189-1
24	36.8	2.3	7811	2	US-09-060-836-1
25	36.8	2.3	7811	2	US-09-184-445-1
26	35.8	2.2	1368	4	US-08-972-927-8
27	35.6	2.2	1254	4	US-09-134-001C-973

28	35.4	2.2	8743	3	US-09-081-320-1	Sequence 1, Appl1
29	35.4	2.2	8743	4	US-09-574-141A-1	Sequence 1, Appl1
30	35.4	2.2	8743	4	US-09-707-780-1	Sequence 1, Appl1
31	35.2	2.2	2343	4	US-09-641-638-652	Sequence 652, App
32	35.2	2.2	4258	3	US-07-765-830A-5	Sequence 5, Appl1
33	35	2.1	2272	4	US-09-108-857-1	Sequence 1, Appl1
34	35	2.1	2418	4	US-09-285-379-1	Sequence 1, Appl1
35	34.6	2.1	2487	4	US-08-377-292-1	Sequence 1, Appl1
36	34.4	2.1	3747	4	US-09-690-364-17	Sequence 1, Appl1
37	34	2.1	5152	4	US-09-690-364-10	Sequence 1, Appl1
38	33.4	2.0	11340	4	US-08-961-527-147	Sequence 147, App
39	33	2.0	4880	4	US-09-402-929-5	Sequence 10, Appl1
40	32.8	2.0	1119	4	US-08-987-943-2	Sequence 2, Appl1
41	32.8	2.0	1161	4	US-09-134-001C-2770	Sequence 2770, Ap
42	32.8	2.0	1755	4	US-08-987-943-1	Sequence 1, Appl1
43	32.8	2.0	2520	1	US-08-087-016-1	Sequence 1, Appl1
44	32.8	2.0	3465	4	US-08-914-999-5	Sequence 5, Appl1
45	32.8	2.0	14176	1	US-08-307-499-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

US-08-828-922-2

Sequence 2, Application US/08828922

Patent No. 5834240

GENERAL INFORMATION:

APPLICANT: Olga, Bandman

APPLICANT: Pirelli, Lal

TITLE OF INVENTION: TRANSFORMING GROWTH FACTOR-B RECEPTOR

TITLE OF INVENTION: ASSOCIATED PROTEIN

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/828, 922

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0258 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1820 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: BRATTU01

CLONE: 753423

US-08-828-922-2

Query Match

Best Local Similarity

62.0%; Score 1011.4; DB 2; Length 1820;

83.3%; Pred. No. 4.7e-299;

RESULT 4
US-09-643-597-36
: Sequence 36, Application US/0966359
: Patent No. 6426072
: GENERAL INFORMATION:
: APPLICANT: Wang, Tongtong
: APPLICANT: Fan, Liqun

Query Match	14.8%	Score 240.8	DB 4	Length 686
Best Local Similarity	79.9%	Pred. No. 1,2e-63		
Matches 362	0	Mismatches 83	Indels 8	Gaps 7

RESULT 5
US-08-232-463-14/c
; Sequence 14, Application US/08232463

```

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTY: USA  
ZIP: 22313-0299  
  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpt-F15  
US-08-232-463-14  
  
Query Match 3.1%; Score 50.6; DB 1; Length 7218;  
Best Local Similarity 5.6%; Pred No. 7,1e-05;  
Matches 23; Conservative 215; Mismatches 170; Indels 0; Gaps 0
```


Db	388	TGCAGACCTCTGGATAGCCCTGGACTTTGGCGCTTTTCCTCGATATCCAGATATCTGGCCAC	447
OY	805	GGCTGGAGAGACTTTAACTACTGTCACAAGTATGATTATTAACAGTGGAGAGAGATTAGATC	864
Db	448	AGGAACTCATGTGCGGAAAAGTAAACATTTTTGGTGTGGAGAACTGGGAAAAAGAAATATTC	507
OY	865	CTACAAAGCTCACTTTGGTCCCATCTTCATCTGTGGATTCACCTCCGTGGATGGGAACTCTA	924
Db	508	TTTGGACACAGAGAGAAAATTCATCTTTACTATTTGCATATATGCTGTGGTGGAAATACCT	567
OY	925	TGCGACGGTTCTGAAGATGGGACATTGAGATTGTGGCAAACTCTGTGTAGAAA	978
Db	568	AGCCAGTGGAGCCATAGATGAATCATCATATTTTTTGGATATTTGCAACTGAAA	621

```

1 RESULT 12
2 US-08-283-917-8
3 ; Sequence 8, Application US/08283917
4 ; Patent No. 5849557
5 ;
6 ; GENERAL INFORMATION:
7 ;
8 ; APPLICANT: ADACHI, HIDEKI
9 ; APPLICANT: TSUJIMOTO, MASAFUMI
10 ; APPLICANT: INOUE, KEIZO
11 ; APPLICANT: ARAI, HIROYUKI
12 ;
13 ; TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
14 ;
15 ; TITLE OF INVENTION: AND GENE THEREOF
16 ;
17 ; NUMBER OF SEQUENCES: 31
18 ;
19 ; CORRESPONDENCE ADDRESS:
20 ;
21 ; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER &
22 ;
23 ; STREET: 1755 S. Jefferson Davis Highway, Suite 400
24 ;
25 ; CITY: Arlington
26 ;
27 ; STATE: Virginia
28 ;
29 ; COUNTRY: U.S.A.
30 ;
31 ; ZIP: 22202
32 ;
33 ; COMPUTER READABLE FORM:
34 ;
35 ; MEDIUM TYPE: Floppy disk
36 ;
37 ; COMPUTER: IBM PC compatible
38 ;
39 ; OPERATING SYSTEM: PC-DOS/MS-DOS
40 ;
41 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
42 ;
43 ; CURRENT APPLICATION DATA:
44 ;
45 ; APPLICATION NUMBER: US/08/283,917
46 ;
47 ; FILING DATE: 03-AUG-1994
48 ;
49 ; CLASSIFICATION: 435
50 ;
51 ; PRIOR APPLICATION DATA:
52 ;
53 ; APPLICATION NUMBER: JP 209943/1993
54 ;
55 ; FILING DATE: 03-AUG-1993
56 ;
57 ; ATTORNEY/AGENT INFORMATION:
58 ;
59 ; NAME: Obion, No. 5849557man F.
60 ;
61 ; REGISTRATION NUMBER: 24,618
62 ;
63 ; REFERENCE/DOCKET NUMBER: 2292-030-0
64 ;
65 ; TELECOMMUNICATION INFORMATION:
66 ;
67 ; TELEPHONE: (703) 413-3000
68 ;
69 ; TELEFAX: (703) 413-2220
70 ;
71 ; TELEX: 248855 OPAT UR
72 ;
73 ; INFORMATION FOR SEQ ID NO: 8:
74 ;
75 ; SEQUENCE CHARACTERISTICS:
76 ;
77 ; LENGTH: 2085 base pairs
78 ;
79 ; TYPE: nucleic acid
80 ;
81 ; STRANDEDNESS: unknown
82 ;
83 ; TOPOLOGY: unknown
84 ;
85 ; MOLECULE TYPE: cDNA
86 ;
87 ; ORIGINAL SOURCE:
88 ;
89 ; ORGANISM: Bos taurus
90 ;
91 ; FEATURE:
92 ;
93 ; NAME/KEY: CDS
94 ;
95 ; LOCATION: 844..2073
96 ;
97 ; US-08-283-917-8

```

Query Match	2.48;	Score 39.6;	DB 2;	Length 2085;
Best Local Similarity	47.98;	Pred. No. 0.076;		
Matches 145;	Conservative	0;	Mismatches 155;	Indels 3;
				Gaps 1;

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: TWO NEW WD-40 PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,534
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0332 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2003, 10:00:47 ; Search time 2440 Seconds

(without alignments)
10819.116 Million cell updates/sec

Title: US-09-856-836-1

Sequence: 1 ttaccgcctaccgcgtggga.....atgcctctaataaaaaa 1630

Scoring table: IDENTITY_MNC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbu:*
5: em_estbv:*
6: em_estbl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1009.6	61.9	1860	11	AF161496 Homo sapi
2	872	53.5	933	14	BQ936632 AGENCOURT
3	819	50.2	948	13	B1407527 602918767
4	720.8	44.2	904	9	AU123736 AU123736
5	703.8	43.2	752	12	BF023572 BF023572
6	667	40.9	944	9	AL517952 AL517952

7	658.8	40.4	684	9	AV060640
8	658	40.4	880	14	BO219033
9	651.8	40.0	702	12	BF720733
10	648.8	39.8	1174	14	BM802585
11	645.8	39.6	1058	12	BF682981
12	635.2	39.0	902	14	BQ889506
13	632.4	38.8	926	13	BI687709
14	631.6	38.7	875	14	BQ439153
15	631	38.7	1005	13	BI408095
16	629.8	38.6	1055	13	BM472016
17	629.4	38.6	839	12	BQ066448
18	624.6	38.3	656	13	BI112632
19	614.2	37.7	1019	13	BM450309
20	612.2	37.6	800	9	AU128959
21	611.6	37.5	901	9	AL518570
22	609	37.4	881	9	AU132581
23	608.2	37.3	757	12	BG687473
24	607.4	37.3	632	10	AM744196
25	605	37.1	620	12	BF467887
26	603.6	37.0	901	9	AL556515
27	599	36.7	1086	14	BQ433088
28	597.6	36.7	879	14	BQ222878
29	596.8	36.6	673	12	BF723058
30	588.6	36.1	1038	9	AL545602
31	585.2	35.9	894	12	BF118711
32	580.4	35.6	874	9	AL570903
33	579.8	35.6	885	14	BQ883467
34	577.6	35.4	955	14	BQ672095
35	576.4	35.4	833	9	AU119814
36	573	35.2	796	12	BG778828
37	571.6	35.1	1093	9	AL545569
38	570	35.0	959	9	AL547128
39	569.8	35.0	962	9	AL539156
40	569.4	34.9	704	12	BG740119
41	568.6	34.9	951	9	AL551950
42	563.4	34.6	629	10	AM682227
43	562.2	34.5	973	9	AL547220
44	561	34.4	664	13	BI596603
45	560	34.4	952	9	AL539237

ALIGNMENTS

RESULT 1	AF161496	1860 bp	mRNA	linear	HTC 22-MAY-2001
LOCUS	AF161496				
DEFINITION	Homo sapiens HSPC147 mRNA, complete cds.				
ACCESSION	AF161496				
VERSION	AF161496.1	GI:6841515			
KEYWORDS	HTC.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 1860) Zhang, Q.H., Ye, M., Wu, X.Y., Ren, S.X., Zhao, M., Zhao, C.J., Fu, G., Shen, Y., Fan, H.Y., Lu, G., Zhong, M., Xu, X.R., Han, Z.G., Zhang, W., Tao, J., Huang, Q.H., Zhou, J., Hu, G.X., Gu, J., Chen, S.J., and Chen, Z.				
TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells				
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)				
MEDLINE	20499367				
PUBMED	11042152				
REFERENCE	2 (bases 1 to 1860)				
AUTHORS	Ye, M., Zhang, Q.H., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J., and Chen, Z.				
TITLE	Human full length cDNA cloned from cd34+ stem cells				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 1860)				
AUTHORS	Ye, M., Zhang, Q.H., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J., and Chen, Z.				

strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - Oligo(dT) primer [5', TGTTACCAATCTCAGAGGAGGAGGCGCCCTGTTTATTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 261 a 237 c 176 g 274 t
ORIGIN

Query Match 50.2%; Score 819; DB 13; Length 948;
Best Local Similarity 97.0%; Pred. No. 1.9e-211;
Matches 919; Conservative 0; Mismatches 17; Indels 11; Gaps 8;

```

OY 633 ATATGCTGTAGCAGCATGGATATATCTCTGAAGCAGATTTGGTTATTTACTTATG 692
DB 947 AATGCTCTTGAGCAGCATGGATATATCTCTGAAGCAGATTTGGTTATTTACTTATG 888
OY 693 GAGCATCTATTTGCTTTCATATGAGTAACT--CTGAGGCCAATTAATCCTTTGAAGC 750
DB 887 GAGCATCTA-TGCTTTTCATATGCTGAGTAACTTTCTGGAGCCCAATTAATCCTTTGAAGC 829
OY 751 TCTGCGACATCAATTTCTGCTC-TTTTCATCCAGAGAGAGTTTCTTGTGCGGTG 809
DB 828 TCTGCGACATCAATTTCTGCTCCTTTTCATCCAGAGAGAGTTTCTTGTGCGGTG 769
OY 810 GAGAGAC-TTTAACTGTACAGATATATTAACATGAGAGAGTTTCTTGTGCGGTG 867
DB 768 GAGAGACTTTTAACTGTACAGATATATTAACATGAGAGAGTTTCTTGTGCGGTG 709
OY 868 CAAAGTCACTTTGCTTCCATTCATCTGTGAGATTCAGTCTGATGGGGAATCTATGC 927
DB 708 CAAAGTCACTTTGCTTCCATTCATCTGTGAGATTCAGTCTGAT--GGAGACTCTATGC 650
OY 928 CAGCGGTTCTGAAGTGGGACATTTGTTGGCAACTGTGGTGAAGAAAGCTATGG 987
DB 649 CAGCGGTTCTGAAGTGGGACATTTGTTGGCAACTGTGGTGAAGAAAGCTATGG 590
OY 988 CCGTGGAAATGGCGTNTCCGAGAGAGACGCGGGAATCGCAAAAGCAATCGG 1047
DB 589 CCGTGGAAATGGCGTNTCCGAGAGAGACGCGGGAATCGCAAAAGCAATCGG 530
OY 1048 ATTTCAGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1107
DB 529 ATTTCAGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473
OY 1108 CATCTATTCATCACTCTGAGATTAAGGCGTGAAGATCAGACGAGTGTGCGGAACA 1167
DB 472 CATCTATTCATCACTCTGAGATTAAGGCGTGAAGATCAGACGAGTGTGCGGAACA 413
OY 1168 TATGTTCAATGATCAATCAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1227
DB 412 TATGTTCAATGATCAATCAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 353
OY 1228 AGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1287
DB 352 AGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 293
OY 1288 TGAAGTGTCCCGTGAGTGAATATGAGTGTGAGAGTGAAGAGAGAGAGAGAGAG 1347
DB 292 TGAAGTGTCCCGTGAGTGAATATGAGTGTGAGAGTGAAGAGAGAGAGAGAGAG 233
OY 1348 TCACATAGTGCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1407
DB 232 TCACATAGTGCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 173
OY 1408 ACCTAATATTTCTTTTATGCTTTATATTTATGAAGAAAGAAATATTTGGCCATTTTTC 1467
DB 172 ACCTAATATTTCTTTTATGCTTTTATATTTATGAAGAAAGAAATATTTGGCCATTTTTC 113
OY 1468 TCAGCTTCCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1527

```

Db 112 TGACTTCCCTTAAG 53
OY 1528 ACATGATTAAGTAACTACCGGTTGATCT-CTTTCATTTGACAGAGACTG 1573
Db 52 ACATGATTAAGTAACTACCGGTTGATCTCTTTCATTTGACAGAGAGAGACTG 6

RESULT 4

AU123736 904 bp mRNA linear EST 01-AUG-2002
LOCUS AU123736
DEFINITION AU123736 NT2RM2 Homo sapiens cDNA clone NT2RM2000935 5', mRNA

ACCESSION AU123736
VERSION AU123736.1 GI:10948452
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 904)
Ota, T., Nakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and Isogai, T.

TITLE HRI human cDNA project (Ota, T., Nakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S., Isogai, T.)
JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomisehri.co.jp

FEATURES
source
1..904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM2000935"
/clone_id="NT2RM2"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"
Location/Qualifiers

BASE COUNT 257 a 173 c 214 g 256 t 4 others
ORIGIN

Query Match 44.2%; Score 720.8; DB 9; Length 904;
Best Local Similarity 89.3%; Pred. No. 9.7e-185;
Matches 806; Conservative 0; Mismatches 94; Indels 3; Gaps 3;

```

OY 64 CCGCTCCCTCCCTCCCTCCGCGCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
DB 2 CCGCTCCGCTTGGCGCGCGCGCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 61
OY 124 CAGCGCGCGCGTGGTGGATNTGGCTTTCAGCGGATCAGCGCTTACGCTTCTGAT 183
DB 62 CAGCGAGACCGGTGGTGGATNTGGCTTTCAGCGGATCAGCGCTTACGCTTCTGAT 121
OY 184 CAGCGCTTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
DB 122 CAGCGCTTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 181
OY 244 AACATTTTGGGTCATAAAGTCTGTTGGGTGCAACATTGAATAGAGATCCACCAA 303
DB 182 AACATTTTGGGTCATAAAGTCTGTTGGGTGCAACATTGAATAGAGATCCACCAA 241
OY 304 AGCTGGACAGAGCTGCAGACTTCAAGCCAAAGATGAGAGAGAGAGAGATGA 363

```



```

DB 97 GATTAAGTAAACCGTTGATCTCTTTCATTTGTAAGAGCTGTCGAGACAGCTCATAT 38
OY 1592 TTTTATGTTATCTAAATAATGCTCTAAATAATAAAAA 1628
DB 37 TTTTAGTTATCTAAATAATGCTCTAAATAATAATA 1

RESULT 6
AL517952 944 bp mRNA linear EST 13-FEB-2001
LOCUS AL517952
DEFINITION AL517952 ltrL_NFL011.NBC1 Homo sapiens cDNA clone CS0DA004YH09 5
ACCESSION AL517952
VERSION AL517952.1 GI:12781445
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 944)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 944
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DA004YH09"
/clone_1bp="LTR_NFL011.NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-Oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact: Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax: (1) 301 610 8371 Email:
f.liang@lifestech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT 259 a 197 c 235 g 250 t 3 others
ORIGIN
Query Match 40.9%; Score 667; DB 9; Length 944;
Best Local Similarity 89.9%; Pred. No. 4,3e-170;
Matches 712; Conservative 1; Mismatches 79; Indels 0; Gaps 0;

OY 66 TCCCTCCCTCCCTGTCGCCATGCGCATGAGCAGCGCGCTCACTGCTGGGCCACA 125
DB 138 TCGGTGGCTTCCCGCCCGCCATGAGCAATGAGACGCGCTCACTGCTGGGCCACA 197
OY 126 CGCGGCGCGCTGTGATNTGGCTTGAAGGCGATCAGCGCTTACGGCTACTTCTGATCA 185
DB 198 CGCGACCCCTGGTGTGATTTGGCTTCAAGTGCAATCAGCGCTTATGGGTATTTCTAATCA 257
OY 186 GGGCTTGCAAGATGAGCCATGCTCCGCGAGAGATACAGAGACTGATTTGAA 245
DB 258 GGGCTTGCAAGATGAGTGAACATGCTACGCGAGGAGATACAGAGACTGATTTGAA 317
OY 246 CATTTTGGGTCAATAAGGTCTGTTTGGGGTCAACATTGAATAAGATGCCACAAAG 305
DB 318 CATTTTGGGTCAATAAGGTCTGTTTGGGGTCAACATTGAATAAGATGCCACAAAG 377
OY 306 CTGCGACAGAGCTGACAGACTTACAGCCAAAGTATAGGGATGGGCTCAGAGATGAT 365

```

```

DB 378 CAGCTACAGAGCTGACGATTTTCACAGCCAAAGTGTGGAGTGTCTGTCAGAGATGAAT 437
OY 366 TGATGACCCCGGCGCATTAAGACATTTGCAAGACTGTGATTTTCACACAGATGAAT 425
DB 438 TGATGACCCCGGCGCATTAAGACATTTGCAAGACTGTGATTTTCACACAGATGAAT 497
OY 426 ACCTGNTAACTGGGGGACAGAGTAAACTGTCGCGCATATATGACTTGAAACAACTGAG 485
DB 498 ATTGTTTAACCGGGGACAGAGATTAACCTGTACCATATATGACTTGAAACAACTGAG 557
OY 486 CAGAACCTTAAGGAATTCAGTGCGCCACACTTCTGATTAATAAAGCTGTGGTGCAGTG 545
DB 558 CAGAACCTTAAGGAATTTAGTGCTACTTCTGTATATAAAGCTGTGGTGCAGTG 617
OY 546 ACGATTAACAGATCTCTTACGCGATTAACACTGTCGCGCTGCGATGATGCCACAA 605
DB 618 AGATTAACAGATCTCTTACGCGATTAACACTGTCGCGCTGCGATGATGCCACAA 677
OY 606 TGACAGAAAGTGAAATCTCTGAAATTTAATGCTGTAGCAGCATGATATCTCTG 665
DB 678 TGACAGAAAGTGAAATCTCTGAAATTTAATGCTGTAGTATGGAATATATCTCTG 737
OY 666 AAGGAGAGATTTTGGTATTAATCTTATGACATCATGCTTTTCAATGTCAGTACGC 725
DB 738 AGGAGAGATTTTGGTATTAATCTTATGACATCATGCTTTTCAATGTCAGTACGC 797
OY 726 TGACACCAATTAATCTTGAAGCTCTGCGACCATCATATCTCGTCTTTCATCCAG 785
DB 798 TGACACCAATTAATCTTGAAGCTCTGCGACCATCATATCTCGTCTTTCATCCAG 857
OY 786 AGAGAGATTTTCTTGTGCGGCTGAGAGAACTTTAACTGTACAGATTTATTAACA 845
DB 858 ACAAAGAAATTTCTTGTGCGGCTGAGAGAACTTTAACTGTATGAATGATTTAATA 917
OY 846 GTGAGAGAGAGT 857
DB 918 GTGGGAGAGATT 929

RESULT 7
AV060640 684 bp mRNA linear EST 24-OCT-2001
LOCUS AV060640 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
DEFINITION
Accession AV060640.2 GI:16380782
VERSION
KEYWORDS
SOURCE EST.
ORGANISM house mouse.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 684)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

```

REFERENCE: 1 (bases 1 to 1058)
AUTHORS: Nih-MGC
TITLE: National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL: Unpublished (1999)
COMMENT: Contact: Robert Strausberg, Ph.D.
Email: c9agbbs-r@mail.nih.gov
Tissue Procurement: M. Bento Soares, Ph.D.
CDNA Library Preparation: Bento Soares and M. Fatima Bonaldo

QY 246 CATTTTGGGTCATTAAGTGTCTTTGGGGTCAACATTTGAATTAAGATGCCACCAAG 305
 |||||||
 Db 275 CATTTTGGGTCATTAAGTGTCTTTGGGGTCAACATTAAGATGCCACCAAG 334
 |||||||
 QY 306 CTGCGACAGCAGCTGCGACACTTTCACAGCCAAAGTATGGATGCGGTCTCAGAGATGAAT 365
 |||||||
 Db 335 CAGTACAGCAGCTGCGACACTTTCACAGCCAAAGTATGGATGCGGTCTCAGAGATGAAT 394
 |||||||
 QY 366 TGATGACCCCTGGCTCATTAAGCAGATTTGCAAGACTGTGGATTTACACAGAGATGAAT 425
 |||||||
 Db 395 TGATGACCCCTGGCTCATTAAGCAGATTTGCAAGACTGTGGATTTACACAGAGATGAAT 454
 |||||||
 QY 426 ACTGNTAACTGGGGAGCAGATTAACCTGCTGATTAATGATTAACCAAACTGAAG 485
 |||||||
 Db 455 ATTTGTTAACCGGGGAGCAGATTAACCTGTTACGATTAATGATTAACCAAACTGAAG 514
 |||||||
 QY 486 CAGAACCTAAGAAATAGTGGCCACACTTCTGTATTAATAAAGCTCTGTGGTGCAGTG 545
 |||||||
 Db 515 CAGAACCTAAGAAATAGTGGTGCATCTTGTATTAATAAAGCTCTGTGGTGCAGTG 574
 |||||||
 QY 546 ACGATTAACAGATCTTTCAGCGGATTAATAACTGTGGCTGTGGATTCACCAAA 605
 |||||||
 Db 575 AGGATTAACAGATCTTTCAGCGGATTAATAACTGTGGCTGTGGATTCACCAAA 634
 |||||||
 QY 606 TGACAGAGTGAATCTCTGATTTATATATGCTGTAGAGAGATGAGATTAATCTCTG 665
 |||||||
 Db 635 TGACAGAGTGAATCTCTGATTTATATATGCTGTAGAGATGAGATTAATCTCTG 694
 |||||||
 QY 666 AAGGAGAGATTTGGTTATTAATGAGAGATTAATGCTTTTCATAGTGAAGTGC 725
 |||||||
 Db 695 AAGGAGAGATTTGGTTATTAATGAGAGATTAATGCTTTTCATAGTGAAGTGC 754
 |||||||
 QY 726 TGAGCCAAATTAATCTTTCAGAGCTCTGCGACCATTAATGCTCTTTTCATAGTGC 785
 |||||||
 Db 755 TGAGCCAAATTAATCTTTCAGAGCTCTGCGACCATTAATGCTCTTTTCATAGTGC 814
 |||||||
 QY 786 AGAAGAGATTTGTTGTC--GGGTGAGAGAGACTTTAACTGATGATGATTA 840
 |||||||
 Db 815 AGAAGAGATTTGTTGTCAGAGGCGGAGAGATTTAACTGATGATGATTA 871
 |||||||

RESULT 13
 LOCUS B1687709 926 bp mRNA EST 18-SEP-2001
 DEFINITION 603315616f1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:535586 5',
 mRNA sequence.
 ACCESSION B1687709
 VERSION B1687709.1 GI:15650337
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 REFERENCE NIH-MGC http://mgi.mc.man.ac.uk/.
 1 (bases 1 to 926)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs.fda.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLN at:
 http://image.llnl.gov
 Plate: LLM11904 row: o column: 03
 High quality sequence stop: 759.
 FEATURES
 source Location/Qualifiers
 1..926
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"

/clone="IMAGE:535586"
 /clone_1b="NCI_CGAP_Mam6"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-Sport6; Site: 1; Salt:
 library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"
 BASE COUNT 276 a 190 c 199 g 260 t 1 others
 ORIGIN

Query Match 38.8%; Score 632.4; DB 13; Length 926;
 Best Local Similarity 98.3%; Pred. No. 1.1e-160;
 Matches 650; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY 968 GTGGTGAAGAAAGCTATGGCTGTGGAATGCGTGTCTGAGAGAGAGCGGGAA 1027
 |||||||
 Db 1 GTGGTGAAGAAAGCTATGGCTGTGGAATGCGTGTCTGAGAGAGAGCGGGAA 60
 |||||||
 QY 1028 CTGGCAAGCGCAAGATCGGATTTCCAGAAACAGAGAGAGAGCTGGCAGAGAAAT 1087
 |||||||
 Db 61 CTGGCAAGCGCAAGATCGGATTTCCAGAAACAGAGAGAGAGCTGGCAGAGAAAT 116
 |||||||
 QY 1088 GCTTCAGAGATTCAGATTCATCTATTCATCAACCTCGAAGTTAAGCCGAGCATCA 1147
 |||||||
 Db 117 GCTTCAGAGATTCAGATTCATCTATTCATCAACCTCGAAGTTAAGCCGAGCATCA 176
 |||||||
 QY 1148 GAGCTGCTGCCGAACCATATGTCATGAGCTAACAAGAGAGAGAGAGCTCGGCT 1207
 |||||||
 Db 177 GAGCTGCTGCCGAACCATATGTCATGAGCTAACAAGAGAGAGAGAGCTCGGCT 236
 |||||||
 QY 1208 TCAGATTAAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267
 |||||||
 Db 237 TCAGATTAAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 296
 |||||||
 QY 1268 TCAGATTAAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1327
 |||||||
 Db 297 TCAGATTAAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 356
 |||||||
 QY 1328 TGAGGAG 1387
 |||||||
 Db 357 TGAGGAG 416
 |||||||
 QY 1388 ACTTACATCTCCATTTTACACCTTAATTTTCTTTAGCTTTATGATGAAGAGAA 1447
 |||||||
 Db 417 ACTTACATCTCCATTTTACACCTTAATTTTCTTTAGCTTTATGATGAAGAGAA 476
 |||||||
 QY 1448 AATTAATTTGCTATTTTCTGACTTCCCTTAAGAGAGAGAGAGAGAGAGAGAGAG 1507
 |||||||
 Db 477 AATTAATTTGCTATTTTCTGACTTCCCTTAAGAGAGAGAGAGAGAGAGAGAGAG 536
 |||||||
 QY 1508 AGTGAAG 1567
 |||||||
 Db 537 AGTGAAG 596
 |||||||
 QY 1568 GAGCTGCTTCAAG 1627
 |||||||
 Db 597 GAGCTGCTTCAAG 656
 |||||||
 QY 1628 A 1628
 |||||||
 Db 657 A 657
 |||||||

RESULT 14
 LOCUS BQ439153 875 bp mRNA EST 24-MAY-2002
 DEFINITION AGENCOURT_7827244 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014262
 5', mRNA sequence.
 ACCESSION BQ439153
 VERSION BQ439153.1 GI:21178229

ORIGIN

Query Match 38.7%; Score 631; DB 13; Length 1005;
Best Local Similarity 89.6%; Pred. No. 2.8e-160;
Matches 772; Conservative 0; Mismatches 71; Indels 19; Gaps 8;

QY 781 TCCAGAGAAAGAGTTCTTGTGGGGTGGAGAACTTTAACTGTACAGTATGATTA 840
DB 1004 TTTCATCCAGAGAAAGATTTCTTTTGGGTGGAAAACTTTAACCTGGACAGTATGATTA 945
QY 841 TAACAGTGGAGAA--GAGTTAGATCTCAAAAGGTCACTTTGGTCCATTCACGT--GT 897
DB 944 TAGCAGTGGAGAAAGAGTTGAATCCCTACAGAGTCCCTTTGGTCCATTCACGTGGG 885
QY 898 GAGATTCAGTCTGTGGGGAAGTCTAT--GCCAGCGGTTCTGAAGATGGACATTTGAG 954
DB 884 GAGATTCAGTCTGTGGGGAAGTCTATGCTCAAGCGGTTCTGAAGATGGACATTTG 825
QY 955 A-----TTGTGGCAAACTGTGTAGAGAAAGCTATGGCTGTGGAA--TGGGTGNTT 1006
DB 824 AGATTTGTGGCAAACTGTGTAGAGAAAGCTATGGCTGTGGAAAGTCCGTCTTTC 765
QY 1007 CCTGAGAGAGACAGGGGGAACTGGCAAAAGCCAAAGATCGGATTTCCAGAAACAGCAGAG 1066
DB 764 CCTGAGAGAGACAGGGGGAACTGGCAAAAGCC--AAGATCGGATTTCCAGAAACAGCAGAG 706
QY 1067 GAAGAGTGGCAGAGAAATTTGCTTCAGAGATTCAGATTCATCTATTCACAGTCT 1126
DB 705 GAAGAGTGGCAGAGAA--AATTTGCTTCAGAGATTCAGATTCATCTATTCACAGTCT 648
QY 1127 GAAGTTAAGGCTGAGCATCAGAGTGTGTGCGGAAACCATATGTCATGAGACTAACA 1186
DB 647 GAAGTTAAGGCTGAGCATCAGAGTGTGTGCGGAAACCATATGTCATGAGACTAACA 588
QY 1187 AGCAGAGACAAGCATCCGCTTCAGAGTACTGTCTGCTGAGGCAAAAGAGGCGAGAAA 1246
DB 587 AGCAGAGACAAGCATCCGCTTCAGAGTACTGTCTGCTGAGGCAAAAGAGGCGAGAAA 528
QY 1247 TATTGGGGCATATGATTTAGTTCAGTGCAGAGAAACAGTACTAGTGTGCGGCTGAGTG 1306
DB 527 TATTGGGGCATATGATTTAGTTCAGTGCAGAGAAACAGTACTAGTGTGCGGCTGAGTG 468
QY 1307 AAAATGGCTAGTGTCTGAGGTGCGAGGAGAGGATTTGCTCAGATAGTGCATAGCCT 1366
DB 467 AAAATGGCTAGTGTCTGAGGTGCGAGGAGAGGATTTGCTCAGATAGTGCATAGCCT 408
QY 1367 GCTGTTGGAATGAAAAGCCAACTTACATCTCATTTTACACCTAAATTTCTTTAGCT 1426
DB 407 GCTGTTGGAATGAAAAGCC--CTTACATCTCATTTTACACCTAAATTTCTTTAGCT 350
QY 1427 GTTTATGTTATGAAGAAATATATTGGCCTATTTTTCAGCTTTCCTTAAAGAAAG 1486
DB 349 GTTTATGTTATGAAGAAATATATTGGCCTATTTTTCAGCTTTCCTTAAAGAAAG 290
QY 1487 AATGCCCTTTTGTCTGCTGCTAGTATGAAGAGAGAAATACATGATTAAGTAACCGGT 1546
DB 289 AATGCCCTTTTGTCTGCTGCTAGTATGAAGAGAGAAATACATGATTAAGTAACCGGT 230
QY 1547 TTGATCTCTTTCATTTACAGAGAGTGTTCAGAACAGCTCATATTTTATGTTATCTAAA 1606
DB 229 TTGATCTCTTTCATTTACAGAGAGTGTTCAGAACAGCTCATATTTTATGTTATCTAAA 170
QY 1607 TAAAATGCCCTTAAAATTAATAA 1628
DB 169 TAAAATGCCCTTAAAATTAATAA 148

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2003, 13:05:33 ; Search time 3249 Seconds

(without alignments)
3144.070 Million cell updates/sec

Title: US-09-856-836-2
Sequence: 1 MAMQTPLTCSGHTRPVVDX.....EETASSENSDSITSTPEVKA 351

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cg2_1/USFTO.spool/US0856836/runat.31012003.144237.22195/app_query.fasta.1.519
-DB=GenDb1 -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPT=0 -LOOPTXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsm62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=ptc -NORM=ext -HEATSIZE=500 -MINLEN=2000000000
-USRR=US0856836.ecgn.1.1.1616.ernat.31012003.144237.22195 -NCPU=6 -ICPU=3
-WARR=NO.MMAP -IAREG=OEBRY -NRC_SCORES=0 -WAIT -LONGLOG -DEV_TIMECUT=120
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

GenDb1:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.in:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rtd:*
36: em.htg.mam:*
37: em.htg.vtc:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1832	99.6	1630	10 AF096285	AF096285 Mus muscu
2	1784.5	97.0	1563	9 AY049776	AY049776 Homo sapi
3	1784.5	97.0	1712	9 BC000162	BC000162 Homo sapi
4	1784.5	97.0	1716	9 HA010025	AJ010025 Homo sapi
5	1780.5	96.8	1867	9 HSM801659	AL136691 Homo sapi
6	1779.5	96.7	1820	6 AR053395	AR053395 Sequence
7	1777.5	96.6	1864	9 AB024327	AB024327 Homo sapi
8	1548.5	84.2	154115	9 AC010878	AC010878 Homo sapi
9	1418	77.1	1829	5 AF315726	AF315726 Homo sapi
10	1334.5	72.5	187017	2 AC010101	AC010101 Homo sapi
11	1328.5	72.2	166680	2 AC010132	AC010132 Homo sapi
12	1138	61.3	73038	2 AC020884.3	Continuation (4 of
13	1128	61.3	226520	10 AL645802	AL645802 Mouse DNA
14	1052.5	57.2	170393	10 AY061000	AY061000 Mus muscu
15	1052.5	57.2	1591	3 AC020469	AC020469 Drosophi
16	876.5	47.6	31772	2 AC009257	AC009257 Drosophi
17	876.5	47.6	188272	3 AC005639	AC005639 Drosophi
18	876.5	47.6	218565	3 AE003461	AE003461 Drosophi
19	876.5	47.6	295225	3 AE003461	AE003461 Drosophi
20	781.5	42.5	1411	8 AY084604	AY084604 Arabidops
21	766.5	41.7	1399	8 AY074392	AY074392 Arabidops
22	549.5	29.9	81370	8 AB017071	AB017071 Arabidops
23	533	29.0	413	9 HS056430	HS056430 Human HeLa
24	517.5	28.1	313	6 AX261089	AX261089 Sequence
25	506	27.5	91001	8 F911	AC013453 Arabidops
26	506	27.5	141984	8 F911	AC013453 Arabidops
27	489	26.6	98151	2 AC119404	AC119404 Magnapor
28	487	26.5	311	6 AC193751	AC193751 Sequence
29	477	25.9	34410	3 AC008242	AC008242 Leishmani
30	477	25.9	116000	2 AC129714	AC129714 Leishmani
31	454.5	24.7	89765	2 AC017870	AC017870 Drosophi
32	454.5	24.7	188873	3 AC009009	AC009009 Drosophi
33	454.5	24.7	196337	3 AC005894	AC005894 Drosophi
34	454.5	24.7	242172	3 AE003832	AE003832 Drosophi
35	452	24.6	117737	8 F6D8	AC008016 Arabidops
36	452	23.1	126253	8 AC019018	AC019018 Arabidops
37	425	23.1	686	6 AR144171	AR144171 Sequence
38	425	23.1	686	6 AR176358	AR176358 Sequence
39	425	23.1	686	6 AX365643	AX365643 Sequence
40	394	21.4	10048	3 AY061931	AY061931 Drosophi
41	366.5	19.9	1205	8 AF285835	AF285835 Arabidops
42	366.5	19.9	1205	8 AF370485	AF370485 Arabidops
43	363.5	19.8	1011	8 AY064633	AY064633 Arabidops
44	347.5	18.9	1138	8 ATU36765	ATU36765 Arabidops
45	344.5	18.7	1143	8 AF335551	AF335551 PhaseoIus

RESULT 1

ALIGNMENTS

LOCUS	AF096285	1630 bp	mRNA	linear	ROD 27-DEC-1998
DEFINITION	Mus musculus serine-threonine kinase receptor-associated protein				
ACCESSION	AF096285				
VERSION	AF096285.1				
KEYWORDS	GI:4063382				
ORGANISM	Mus musculus				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1630)				
AUTHORS	Data P.K., Chytil A., Gorska A.E. and Moses H.L.				
TITLE	Identification of STRAP, a novel WD domain protein in transforming growth factor-beta signaling				
JOURNAL	J. Biol. Chem. 273 (52), 34671-34674 (1998)				
MEDLINE	99074230				
PUBMED	9856985				
REFERENCE	2 (bases 1 to 1630)				
AUTHORS	Data P.K.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-OCT-1998) The Vanderbilt Cancer Center, Vanderbilt University School of Medicine, 649 Medical Research Building II, Nashville, TN 37232-6838, USA				
FEATURES	Location/Qualifiers				
SOURCE	1..1630				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	/dev_stage="embryo"				
	86..1141				
	/note="WD domain protein: STRAP; TGF-beta receptor-associated protein"				
	/codon_start=1				
	/product="serine-threonine kinase receptor-associated protein"				
	/protein_id="AAC98300.1"				
	/db_xref="gi:4063383"				
	/translation="MAAMQPTLCGHRPPVYDLAFCGTPYGYLLSACKDQKPMIL /OGDWDWIGTELGHRGAWGATLNKDKAAATPAADPAKWDAVSGEMLTAKKH VKTVPDSDNTLILGQDQKILRIYDLKPEPEPEISGHTSGIKALMCSDDKILV ADKTVRLMDHATMTTEVKSLENNMVSSEMEYIPBEILIVTYSRIARVSAVSEPI SEFEATVINSASLIHKEFELVAGGEDFYLKYDYNVSGEELSKYKHFPIHCVRSPSP GELVAVSGEDGLRLMOTVYVGTGYGLMKCVLPEDSGELAPKIGFPTAEELAEEL ASENDSIYSRPEPKA"				
BASE COUNT	455 a	349 c	406 g	420 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	6,69e-158	Length:	1630		
Score:	1832.00	Matches:	347		
Percent Similarity:	98.86%	Conservative:	0		
Best Local Similarity:	98.86%	Mismatches:	4		
Query Match:	99.57%	Indels:	0		
DB:	10	Gaps:	0		
US-09-856-836-2 (1-351) x AF096285 (1-1630)					
QY	1	MetAlaMeTArgGlnInPrrProLeuTThCySSeGlyHISrHArpProValValaSp**	20		
Db	86	ATGAGCCATGGGAGAGAGCCGCTCACTTGTCTGGGCCACACGGGGCCCGTGGAGATTG	145		
QY	21	AlAPheSeSerGlyIleThrProTyrGlyTyrPheLeuIleSeRaLaCyLSySaSpGlyLyS	40		
Db	146	GCCTTACGGGAGACACAGCCTTACGGCTACTTCTGATCAGCCCTTGCAAAAGATGGCAAG	205		
QY	41	PromELeuArGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyYHISLyGly	60		
Db	206	CCCATGCTCCGCCACGAGATACAGGAGACTGGATTGGAACATTTTGGGTCAATAAGGT	265		
QY	61	AlAlaVTrpGlyAlaTrpLeuAsnLyASpAlaThrLySaLaAlaTrpAlaAlaAlaASP	80		
Db	266	GCTGTGTTGGGGTGCACATTTGAATTAAGAGATGCCACCAAAAGCTGGACAGCAGCTGCAGAC	325		

QY	81	PheThrAlaIysValITrPaspAlaIValSerGlySpGluMetThrLeuAlaHisIys	100
Db	326	TTCCAGGCGAAAGTATGGGATGCGGCTCTCAGAGACATATATTTGATGACCCCTGCTCAATAG	385
QY	101	HisIleValIysThrValAspPheThrGlnAspSerAsnTyrIleu**ThrGlyGlyGln	120
Db	366	CACATTTGTCAGACACTGTGATTTCCACACAGATATGACATTAACCTGCTACATGGGGGACAG	445
QY	121	AspIysIleuLeuAngIleTyrAspLeuAsnIysProGluIaGluProIysGluIleSer	140
Db	446	GATTAACGCGCGCATATATGACTTGACAAACCTGAGCAGAACTTAAGGAATACGT	505
QY	141	GlyHisThrSerGlyIleLysLysAlaIleuTyrPcysSerAspIysGluIleLeuSer	160
Db	506	GGCCACACTTCGTGGATTTAAAAAGGCTCTGGTGACATGACGATTAACAGATCCCTTTCA	565
QY	161	AlaAspAspIysThrValArgLeuTrpAspHisAlaThrMetCThrGluValIysSerLeu	180
Db	566	CGGATATGATTAACAGTTCGGCTCTGGGATATATGCCAATAGACAGAAATGTAATTCCTG	625
QY	181	AsnPheAsnMetSerValSerSerMetGluTyrIleProGluGluGluIleLeuValIle	200
Db	626	AATTTTAATATCTCTGTAGCAGCATATGAGATATATCTCGAAGGAGATTTTGGTTATT	685
QY	201	ThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe	220
Db	686	ACTTATAGCAGCATTCATTCCTTTCATGTCGAGTACGACAGCCCAATTAATTCCTTT	745
QY	221	GluAlaProAlaThrIleAsnSerAlaSer**HisProGluIysGluPheLeuValAla	240
Db	746	GAACTCCCTGCACCATCATTCCTGCTCTCTTATCCAGAAAGGAACTTTCTTGTTGCG	805
QY	241	GlyGlyGluAspPheIysLeuTyrIysTyrAspTyrAsnSerGlyGluGluLeuGluSer	260
Db	806	GGTGGAGAAAGACTTTAACTGTACAAAGTATGATTTTAAACAGTGGAGAGACTTGTGAAATCC	865
QY	261	TyrLysGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyGluLeuTyr	280
Db	866	TACAAAGTGCACCTTGTGCTCCATTCACCTGTGCAGATTACGTCCTGATGGGGAATCTCAT	925
QY	281	AlaSerGlySerGluAspGlyThrLeuArgLeuTyrPglThrValValGlyLysThrTyr	300
Db	926	GCCAGGCGTTCGAATATGGGACATATGAAATTTGTGGCAAACTGTGTAGGAAAGACCTAT	985
QY	301	GlyLeuTrpLysGlyVal**ProGluGluAspSerGlyGluLeuAlaLysProLysIle	320
Db	986	GGCTGTGGAAATGCGCTGCTCTCGAGGAAGACAGCGGGGAACTGGCAAAAGCAAAATGC	1045
QY	321	GlyPheProGluThrAlaGluGluGluLeuAlaGluGluIleAlaSerGluAsnSerAsp	340
Db	1046	GGATTTCCAGAAACACAGACAGAGGAAGCGTGGCAGAAGAATTCCTTCAGAGATTCAGAT	1105
QY	341	SerIleTyrSerSerThrProGluValIysAla	351
Db	1106	TCCATCTATTCATCACTCTCTGAAATTAAGGCC	1138
RESULT 2			
LOCUS	AY049776	1563 bp	mRNA linear PRI 23-OCT-2001
DEFINITION	Homo sapiens unr-interacting protein (UNRIP) mRNA, complete cds.		
ACCESSION	AY049776		
VERSION	AY049776.1	GI:16356636	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	Li, J., Zhou, Y., Peng, X., Yuan, J., and Qiang, B.		
AUTHORS	Submitted (31-JUL-2001) Department of Biochemistry, Institute of Basic Medical Science, Chinese Academy of Medical Sciences, 5 Dong Dan San Tiao, Beijing 100005, P.R. China		
TITLE	Submitted (31-JUL-2001) Department of Biochemistry, Institute of Basic Medical Science, Chinese Academy of Medical Sciences, 5 Dong Dan San Tiao, Beijing 100005, P.R. China		
JOURNAL			

	D	b	648	ACTGATAGACAGCTATTTCGTTCATCAGTCCAGTTAAGTTGGACCCTTAATAATCCTTT 	:	707
Oy	221	GluAlarproalathrilleasnSeraIaSeriHisProgluiysgiUphelueVala	240			
D	708	GAACCTCTCTCAACCATCAATTCGCATCTCTTCATCTCGTAGAAAGAATTTCTTGTTCCA 	767			
Oy	241	GlyGIYgIUaspphelusleutyrlvsttAspyrrtsnsbergiylguileugluer	260			
D	768	GGCGGTACAARITTTAACTTTATMASTATGATTATATATGTGGAGAAATTTCTTGTTGCCA 	827			
Oy	261	TyrILysGLYhisPheglyProIIeHIScSVaIArpPheserProaspGIYgluleutyr	280			
D	828	TACAAGGGACACTTTGGTGCTTACTATTCAGTGTGAGATTTAGTCTGTATGGACAACTTAT 	887			
Oy	281	AIsSerGIYSergIUaspgJYNhleArLeutTrpsInThrVaIVaIGLYsthTYR	300			
D	888	GCCAGTGGTTCACAAAGATGGAGACATTTGGCAAACGTGTGTGAGAAAAAACGTAT 	947			
Oy	301	GIyleutTPlysCySAval**ProGUgluasSergiYGluLenuAIaySPolySIIE	320			
D	948	GGCGTTTGAAAATGTGTGCTCTCTGAAGAAGATATGTGTGAGCTGGCAAGCCAAAGATT 	1007			
Oy	321	GlyPhroGlunhrzlaGuIgUGluLeuaIagLUglUIleaLasergIUasNSerasp	340			
D	1008	GGTTTTCCAGAGCAACAGAGAGGAGCA--GAAGAAATTCCTTCAGAGAAATTCAGAT 	1064			
Oy	341	SerILEtyrsSerThrpProglUVaIVaLSALA 351 ::: :::				
D	1065	TGCATCTTTCCTCACGCTCTCATGTITAAGCC 1097 ::: :::				
RESULT 3						
Locus	BC000162	1712 bp mRNA linear Pri 12-JUL-2001				
DEFINITION	Homo sapiens, un-interrupting protein, clone MGC:5236					
Accession	IMAGE:5900954, mRNA, complete cds.					
Version	BC000162					
Keywords	BC000162.2 GI:14198026					
Source	MGC.					
Organism	Homo sapiens. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
Reference Authors Title Journal	Direct Submission Submitted (03-NOV-2000) National Institutes of Health, Mamalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
REMARK COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov On May 25, 2001 this sequence version replaced gi:12652818. Contact: MGC help desk Email: cgapbs@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA library Arrayed by: The I.M.A.G.E Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hpsc.bcm.tmc.edu/cdna/ Contact: villalon@bcm.tmc.edu. Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Munzy,D.M., Gibbs,R.A.					
FEATURES source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E Consortium/LNLT at: http://image.llnl.gov Series: IRAK Plate: 3 Row: e Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6807652. location/Qualifiers 1..1712					

/translation="MAMROPLTCSGHRPVVDFASGIRPYGELISACKDGRMLR
 OGDGDMTGTFLGKAGVAMGATPLKNDKAKATAADPAKWDVSGDEMLTAKHL
 VKTDFPDSDNATLTLGGDDKLRITDMLKPEAREKESIGRTSGIKALMCSKDQIIS
 ADEKATRLMDHATMEKSLNFMNSVSMEXIIPGELIVYIGSIFHSAVSDPIK
 SEAPATVNSASLHPEKEFLVAGEDYEDYVDSNGELESYGHGPHICVPSPD
 GELVAGSDESDTLRMQTVVGTGTGLMWCVLPEDSDGLAPKIGPETTBELEELA
 SENSDFIPSPDVKA"

BASE COUNT 519 a 320 c 395 g 482 t

ORIGIN

Alignment Scores:

Pred. No.: 1,54e-153 Length: 1716

Score: 1784.50 Matches: 338

Percent Similarity: 97.44% Conservative: 4

Best Local Similarity: 96.30% Mismatches: 8

Query Match: 96.98% Indels: 1

DB: 9 Gaps: 1

US-09-856-836-2 (1-351) x HAM10025 (1-1716)

QY 1 MetalameTarqInThProleuthrCysSerGlyHisThrArgProValasp*** 20
 DB 148 ATGGCAATGAGACAGACCCCTCAGCTCTGACACAGCAGCAGCCGTTGATTG 207
 QY 21 AlaheSerGlyIleThProtyrGlyTrpPheLeuIleSerAlaCysLysAspGlyLys 40
 DB 208 GCCTTCAGTGGCAGCAGCCTTATGGGTAATTCTTAACACGGCTTGCAAGATGCTAA 267
 QY 41 ProMetLeuAtrgInglyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
 DB 268 CCAATGCTACCCAGGAGATACAGGACACGATGGAACATTTTGGCTCATTAAGGT 327
 QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaIleThrAlaAlaAsp 80
 DB 328 GCTGTTGGGGTGAACACTGAATAGATGCCACCAAGCAGCTACAGCAGCTGACAT 387
 QY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
 DB 388 TTCACACCAAAAGTGGGATGCTGCTCAGAGATGAATGATGATGACCTGGCTCATAAA 447
 QY 101 HisLleValLysThrValAspPheThrGlnAspSerAsnTrpLeu***ThrGlyGln 120
 DB 448 CACATTTGCAAGACTGTCGATTCACCAAGATAGTATATTTTGAACCGGGGACAG 507
 QY 121 AspLysLeuLeuArgIleTrpAspLeuAsnLysProGluAlaLysProLysGluLysSer 140
 DB 508 GATTAACCTGTAGCATATATGACTGACAAACCTGACAGCAGACACTAGAAATTAAGT 567
 QY 141 GlyHisThrSerGlyIleLysLysAlaLeuTrpCysSerAspLysGlnIleLeuSer 160
 DB 568 GGTCTATCTCTGGTATATAAAAGCTCTGTGTGACGTGAGATTAACAGATTTCTTCT 627
 QY 161 AlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeu 180
 DB 628 GCGATGACAAACCTGTGACCTTTGGATCAGCTACTGACAGAGAGAAATCTCTA 687
 QY 181 AspPheAsnMetSerValSerSerMetGluTrpIleProGluLysLysLeuValIle 200
 DB 688 AATTTTAATATGCTGTAGTATGATGATATATTCCTGAGGAGAGATTTGTTATTA 747
 QY 201 ThrTyrglyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe 220
 DB 748 ACTTATGAGAGATCTATGCTTTTTCATAGGACAGTAAAGTTTGACCAATTAATCTT 807
 QY 221 GluAlaProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuValAla 240
 DB 808 GAAGCTCTGCAACATCAATCTGCATCTTCATCTGAGAAACAAATTTCTGTTGCA 867
 QY 241 GlyGluAspPheLysLeuTrpLysTrpAspTrpLysSerGlyLysLysLeuLysSer 260
 DB 868 GCGGCGGAGAAATTTTAACCTTATTAAGTATGATTAATATGAGAGAAATTAAGATCC 927
 QY 261 TyrLysGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyLysLeuTrp 280

DB 928 TACAAGGACACTTGGCTATTACAGTGTGAGATTGATGAGAACTTAT 987
 QY 281 AlaSerGlySerGluAspGlyThrLeuArgLeuTrpGlnThrValValGlyLysThrTyr 300
 DB 988 GCCAGTGGTTCAGAAAGATGCAACATTGAGACTAAGCAGAACTGCTGAGAAAAACCTAT 1047
 QY 301 GlyLeuTrpLysCysVal**ProGluLysAspSerGlyLysLeuAlaLysProLysIle 320
 DB 1048 GGCCTTTGGAATATGTGCTCTGAAAGAAATAGTGTAGAGCTGGCAAGCCAAAGATT 1107
 QY 321 GlyPheProGluThrAlaGluGluGluLysLeuAlaGluLysIleAlaSerGluAsnSerAsp 340
 DB 1108 GGTTCACAGACACACAGAAAGAGAGACTA---GAAGAAATGCTTCAGAGAAATTCAGAT 1164
 QY 341 SerIleTrpSerSerThrProGluValLysAla 351
 DB 1165 TGCACTCTTCCTTCAGCTCCTGATGTTAAAGCC 1197

RESULT 5
 HSM801659
 LOCUS
 DEFINITION
 complete cds.
 accession
 version
 keywords
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 1867)
 Wiemann,S., Well,B., Wellenreuther,R., Gassenhuber,J., Glasl,S.,
 Alenorge,W., Boecker,M., Bloeker,H., Baerach,S., Blum,H.,
 Lauber,J., Duesterhoeft,A., Beyer,A., Koehrer,K., Strack,N.,
 Mewes,H.W., Ottenwaelder,B., Obermaier,B., Tampe,U., Heubner,D.,
 Toward a Catalog of Human Genes and Proteins: Sequencing and
 Analysis of 500 Novel Complete Protein Coding Human cDNAs
 Genome Res. 11 (3), 422-435 (2001)
 11230166
 2 (bases 1 to 1867)
 Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
 Wiemann,S.
 Direct Submision
 Submitted (15-JAN-2000) MIPS, Am Klopfersplitz 18a, D-82152
 Martinsried, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by MedGenomix (Martinsried/Germany) within the cDNA
 sequencing consortium of the German Genome Project. This clone
 (DKFZp564N1778) is available at the RZPD in Berlin. Please contact
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES
 source
 location/Qualifiers
 1..1867
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp564N1778"
 /tissue_type="brain"
 /clone_lib="564 (synonym: hfbz2). Vector pAMP1; host
 X1-2blue; sites NotI + SalI"
 /dev_stage="fetal"
 297..1349
 /gene="DKFZp564N1778"
 297..1349
 /gene="DKFZp564N1778"
 /note="unr-interacting protein"
 /product="hypoetical protein"
 /protein_id="CAB6626.1"

CDs
 gene
 CDS

```

/db_xref="GI:6907653"
/translation="MAMROTPLTSCGHTTRPVVDLARSITPYGYFLISACKDKPMIR
OGDTGMIGFLHKGAWGATLNDKATKAATAAFATAVMDVSDMLTAHKHI
VKYVDPTGDSNYLNGGOKLRIDYLNKEAPKESITSGIKRALMCSEROLIS
ADKDYATLMDHATMEYKSLNFMYSMSMEYTEGELIVTVYRSIKALFSAVSDPIK
SEBPATINSALHPEKELVAGEDPKLYKYINSESELESTKGNHGP IHCYREFSD
GELYASGSEDTLRIMQYVGYKLYGLMKCVLPEDSEELAKPKIPEPTEEBLEELA
SENSECIFFSPAPVKA"
polya_signal
1829..1834
polya_site
1842
BASE COUNT 532 a 388 c 422 g 525 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 3,95e-153 Length: 1867
Score: 1780.50 Matches: 337
Percent Similarity: 97.44% Conservative: 5
Best Local Similarity: 96.01% Mismatches: 8
Query Match: 96.77% Indels: 1
DB: Gaps: 1

```

US-09-856-836-2 (1-351) x HSM801659 (1-1867)

```

QY 1 MetAlaMetArgInThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
DB 297 ATGGCAATGAGACAGACCGCGCTGCTGCTGCGCACGCGACCGCTGGTTGATTGG 356
QY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 357 GCCTTCATGGCATCACCCTTATGGTATTTCTTAATCAGCGCTGCAGAAAGATGTGAAA 416
QY 41 PrometLeuArgInGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
DB 417 CCTATGCTACGCGCAGGAGATACAGAGACCTGATTGGAACATTTTGGGTATTAAGGT 476
QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
DB 477 GCTGTTGGGGGCAACACTGATAAGATGCGCAACCAAGCAAGCTACAGCAGCTGCAGAT 536
QY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 537 TTCACAGCAAAAGTGGGATGCTGCTCAGAGATGAATGATGACCTGGCTCATTA 596
QY 101 HisIleValLysThrValAspPheThrGlnAspSerAspTyrLeu***ThrGlyGlyGln 120
DB 597 CACATGTCAGACAGCTGATTTACGACGAGATAGTAATATTGTTTAAACCGGGGACAG 656
QY 121 AspLysLeuLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluIleSer 140
DB 657 GATTAACCTGTACGATATATGACTGAAACCACTGAAGCAGAACCTTAAGGAATTAAT 716
QY 141 GlyHisThrSerGlyIleLysLysAlaLeuTrpCysSerAspAspLysGlnIleLeuSer 160
DB 717 GGTCACTACTCTGGATATAAAAAACCTGTGTCAGTGAGATTAACAGATTTCTTTCT 776
QY 161 AlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeu 180
DB 777 GCTGAGACAAAACCTGTGACTTGGATTCATGCTACTATGACAGAAAGTAATCTCTA 836
QY 181 AsnPheAsnMetSerValSerSerMetGluTyrIleProGluGlyIleLeuValIle 200
DB 837 AATTTAATATGCTGTGTAGTATGAAATATATCTCGAGGGAAGATTTGGTTATA 896
QY 201 ThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe 220
DB 897 ACTTATGAGAGATCTATGCTTTTCAATAGTCGTAAGTTTGGACCAATTAATCTTT 956
QY 221 GluAlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAla 240
DB 957 GAAGGCTCGGCAACATTCATCTCTCATCCCTGAGAAAGAAATTTCTGTTGCA 1016
QY 241 GlyGlyGluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluGluLeuGluSer 260

```

```

DB 1017 GCGCGTGAAGATTTTAAACCTTATAGATGATTATATAGTGAAGAAATAGAACCC 1076
QY 261 TyrLysGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyGluLeuTyr 280
DB 1077 TACAAGGAGACCTTTGGCTTATTCACCTGTGTGAGATTTACTCTGATGAGAACTCTAT 1136
QY 281 AlaSerGlySerGluAspGlyThrLeuArgLeuTrpGlnThrValAlaGlyLysThrTyr 300
DB 1137 GCGATGGTTCAGAGATGAGACATGAGACTTGGCAAACTGTGTGAGAAAAAGCTAT 1196
QY 301 GlyLeuTrpLysCysVal***ProGluLysAspSerGlyGluLeuAlaLysProLysIle 320
DB 1197 GGCCTTGGAAATGTCTGCTCCCTGAGAAAGATAGTGTGAGCTGGCAAGCAAGATTT 1256
QY 321 GlyPheProGluThrAlaGluGluGluLeuAlaGluGluIleAlaSerGluAsnSerAsp 340
DB 1257 GGTTCACAGACACAAAGAGAGAGCTA--GAAGAAATGGCTTCAGAGATTCAGAG 1313
QY 341 SerIleTyrSerSerThrProGluValLysAla 351
DB 1314 TGCAATCTTCCCTTCAGCTCCTGATGTTAAGGCC 1346

```

```

RESULT 6
AR053395 1820 bp DNA linear PAT 29-SEP-1999
LOCUS AR053395
DEFINITION Sequence 2 from patent US 5834240.
ACCESSION AR053395
VERSION AR053395.1 GI:5978257
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1820)
AUTHORS Bandman O. and Lal P.
TITLE DNA encoding a transforming growth factor-.beta. receptor
associated protein
JOURNAL Patent: US 5834240-A 2 10-NOV-1996;
FEATURES
location/Qualifiers
1..1820
source
/organism="unknown"

```

```

BASE COUNT 516 a 370 c 411 g 523 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 4.73e-153 Length: 1820
Score: 1779.50 Matches: 337
Percent Similarity: 97.15% Conservative: 4
Best Local Similarity: 96.01% Mismatches: 9
Query Match: 96.71% Indels: 1
DB: Gaps: 1

```

US-09-856-836-2 (1-351) x AR053395 (1-1820)

```

QY 1 MetAlaMetArgInThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
DB 259 ATGGAATGAGACAGACCGCGCTGCTGCTGCGCACGCGACCGGTGATTG 318
QY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 319 GCTTCAGTGCAGACAGCGCTTATGGTATTTCTTAATCAGCCTTGCAAGATGTGAAA 378
QY 41 PrometLeuArgInGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
DB 379 CCTATGCTACGCGCAGGAGATACAGAGACTGTGGAACATTTTGGGTCAATTAAGGT 438
QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
DB 439 GCTGTTGGGGTGCACACTGAAATGAAGATGCCCAACACAGCTACAGCAGCTGCAGAT 498
QY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 499 TTCACAGCAAAAGTGGGATGCTGCTCAGAGAGATGAATGATGACCCGTGCTCATTA 558

```


QY 101 HistIleValIsthrValAspPheThrGlnAspSerAsnTyrLeu***ThrGlyGlyGln 120
 Db 559 CACATTTGCAAGACTGCTGATTTCCACCAGATGATATTTATTGTTAACCGGGGACAG 618
 QY 121 AspIleLeuAsgIleTyrAspLeuAsnTyrProGluAgluProlGlyIleSer 140
 Db 619 GATTAACCTGTTACCATATATGACTTGAACCAACCTTAAGCAGAACTTAAGAAATTAGT 678
 QY 141 GlyHisThrSerGlyIleTyrLysAlaLeuTyrPcysSerAspAspIleLeuSer 160
 Db 679 GGTCAATCTCTGCTATATAAAAGCTCTGTGGTGCAGTGAAGATTAACAGATTTCTTCT 738
 QY 161 AlaAspAspIsthrValAgluLeuTyrAspHisAlaThrMetThrGluValLysSerLeu 180
 Db 739 GCGATACAAAACCTGTTGCACTTTGGATCAGCTACTATGACAGAACTGAATCTCTA 798
 QY 181 AsnPheAsnMetSerValSerSerMetGluTyrIleProGluGluIleLeuValIle 200
 Db 799 AATTTTAAATGCTGCTGTTAGTATGATGATATATTTCTTGAAGGAGATTTGGTTATA 858
 QY 201 ThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProlIleLysSerPhe 220
 Db 859 ACTATGAGAGATCTATGCTTTTCATAGTGAAGTTGAGCCCAATTAATCTTTT 918
 QY 221 GluAlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAla 240
 Db 919 GAAGCTCTCCACCAACATCAATTCGATCTTCATCTCAGAGAAAGAAATTTCTTTGCA 978
 QY 241 GlyGlyGluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluGluLeuGluSer 260
 Db 979 GGGGGAGAGATTTTAACTTATAGTATGATATATATGATGAGAGATTAAGATTC 1038
 QY 261 TyrLysGlyHisPheGlyProlIleHisCysValAlaArgPheSerProAspGlyLysLeuTyr 280
 Db 1039 TACAAGGACACTTGTCTCTATCTACGTGTGAGATTTAGTCTGATGAGAACTCTAT 1098
 QY 281 AlaSerGlySerGluAspGlyThrIleuArgLeuTyrProGluThrValGlyLysThrTyr 300
 Db 1099 GCGAGTGGTTCACAAAGATGGAACATTTGACATGTCAGCAACCTGCTAGCAAAAGCTAT 1158
 QY 301 GlyLeuTyrLysCysVal***ProGluGluAspSerGlyLysLeuAlaLysProlGly 320
 Db 1159 GGGCTTTGAAATGTGCTCTCTGAAAGATAGTGTGAGCTGCAAAAGCAAGATT 1218
 QY 321 GlyPheProGluThrAlaGluGluLeuAlaGluGluIleAlaSerGluAsnSerAsp 340
 Db 1219 GGTTCCTCAGAGACACAGAAAGAGCTA--GAAGAAATGCTTCAGAGATTCAGAT 1275
 QY 341 SerIleTyrSerSerThrProGluValLysAla 351
 Db 1276 TGCATCTTCTTCAGCTCTGATGTTAAGGCC 1308
 RESULT 7
 AB024327 1864 bp mRNA linear PRI 20-DEC-2000
 LOCUS Homo sapiens pt-wd mRNA for WD-40 repeat protein, complete cds.
 DEFINITION AB024327
 ACCESSION AB024327.1 GI:4519416
 VERSION pt-wd: WD-40 repeat protein.
 KEYWORDS Homo sapiens liver cell_line:HepG2 cDNA to mRNA.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Matsuda, S., Katumata, R., Okuda, T., Yamamoto, T., Miyazaki, K.,
 Senga, T., Machida, K., Thant, A.A., Nakatsugawa, S., and Hamaguchi, M.
 TITLE Molecular cloning and characterization of human MAMD, a novel
 protein containing WD-40 repeats frequently overexpressed in breast
 cancer
 JOURNAL Cancer Res. 60 (1), 13-17 (2000)
 MEDLINE 20110763
 REFERENCE 2 (bases 1 to 1864)

AUTHORS Matsuda, S. and Okuda, T.
 TITLE Direct Submission
 JOURNAL Submitted (25-FEB-1999) Satoru Matsuda, Nagoya University School of
 Medicine, Molecular Pathogenesis; 65 Tsurumai-cho Showa-ku, Nagoya
 466-8550, Japan (E-mail: smatsuda@esu.med.nagoya-u.ac.jp,
 Tel:81-52-744-2463, Fax:81-52-744-2464)
 FEATURES
 source
 1. 1864
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="HepG2"
 /tissue_type="liver"
 1. 1864
 /gene="pt-wd"
 301. 1353
 /gene="pt-wd"
 /function="putative TGF-beta receptor association"
 /note="homologous to TGF-beta-receptor interacting
 protein-1 (TRP-1)"
 /codon_start=1
 /product="WD-40 repeat protein"
 /protein_id="BA07544.1"
 /db_xref="GI:4519417"
 /translation="MAMRQPLTCSGHTRPVVDLAFSDIPYGFILSACKDQPMR
 OGDGDMIGFTLGHKAVMGATLTKDKATKATADATKAMDAVSDELTKLHKHT
 VKTVDFQDSNVLTLTGODKLRITVDLTKPEAPEKEISGHTSKLAKMCSQILS
 ADDKTVRLMDATMTYKSLVFNKNSVSMEITPCEILVITYGKSILAFHSAVSLDPIK
 SFEAPATINSASLHPEKEFEIVAGDEKLYKVDNSGESELYSGHGPPIHCVFSPD
 GELTASGSDTLRLMDQVVGKTVGLMKCVLPERDSELMKPKIGPETEELEETLA
 SENSDCLEPSPDVKA"
 BASE COUNT 528 a 388 c 424 g 524 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.4e-153 Length: 1864
 Score: 1777.50 Matches: 337
 Percent Similarity: 97.15% Conservative: 4
 Best Local Similarity: 96.01% Mismatches: 9
 Query Match: 96.60% Indels: 1
 DB: 9 Gaps: 1
 US-09-856-836-2 (1-351) x AB024327 (1-1864)
 QY 1 MetalameTarGlnThrProLeuThrCysSerGlyHisThrArgProValAsp*** 20
 Db 301 ATGGCAATGAGACAGACCGCTCACCTGCTGCGCCACAGCCACCGTGTGATTG 360
 QY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
 Db 361 GCGTTGATGACATCACCGCTTATGGGTATTTCTTAATCACCGCTGCAAAAGATGTAA 420
 QY 41 PromeTleuArgGlnGlyAspThrGlyAspThrIleGlyThrPheLeuGlyHisLysGly 60
 Db 421 CCTATGCTACGCCAGGAGATACAGAGACTGGATTTGGAACTTTGGGTCATAAAGGT 480
 QY 61 AlaValTyrGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
 Db 481 GCTGTTGGGGTGCACACCTGAATTAAGAGTCCACCAAGAGAGCTACAGAGCTGCAGAT 540
 QY 81 PheThrAlaLysValTyrPaspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
 Db 541 TTCACAGCCAAAGGTGGAGTGTCTCTCAGAGATGAATTAAGTGAACCTGCTCATAAA 600
 QY 101 HistIleValIsthrValAspPheThrGlnAspSerAsnTyrLeu***ThrGlyGlyGln 120
 Db 601 CACATTTGCAAGACTGCTGATTTCCACCAGATGATATTTGTTAACCGGGGACAG 660
 QY 121 AspIleLeuAsgIleTyrAspLeuAsnTyrProGluAgluProlGlyIleSer 140
 Db 661 GATTAACCTGTTACCATATATGACTTGAACCAACCTTAAGCAGAAATTAGT 720
 QY 141 GlyHisThrSerGlyIleTyrLysAlaLeuTyrPcysSerAspAspIleLeuSer 160

repeat_region	/rpt_family="Alu"	1744. .1781	repeat_region	19035. .19080	/rpt_family="ERV1"
repeat_region	/rpt_family="MERL_type"	1793. .2026	repeat_region	19614. .19860	/rpt_family="MIR"
repeat_region	/rpt_family="Alu"	2032. .2132	repeat_region	20803. .20906	/rpt_family="MERL_type"
repeat_region	/rpt_family="MERL_type"	2164. .2252	repeat_region	20994. .21169	/rpt_family="L1"
repeat_region	/rpt_family="L1"	2258. .2613	repeat_region	21214. .21402	/rpt_family="L1"
repeat_region	/rpt_family="L1"	2633. .2895	repeat_region	21437. .21755	/rpt_family="Alu"
repeat_region	/rpt_family="Alu"	2899. .2919	repeat_region	21767. .22143	/rpt_family="L1"
repeat_region	/rpt_family="(TAA)n"	2923. .3480	repeat_region	22168. .22554	/rpt_family="ERV1"
repeat_region	/rpt_family="L1"	3552. .3659	repeat_region	22555. .22619	/rpt_family="MERL"
repeat_region	/rpt_family="L1"	3670. .3777	repeat_region	22620. .22905	/rpt_family="Alu"
repeat_region	/rpt_family="L1"	4243. .4534	repeat_region	22906. .23334	/rpt_family="MERL"
repeat_region	/rpt_family="Alu"	4827. .5132	repeat_region	23335. .23365	/rpt_family="ERV1"
repeat_region	/rpt_family="Alu"	5718. .5925	repeat_region	23835. .24126	
repeat_region	/rpt_family="MERL_type"	6210. .6792			
repeat_region	/rpt_family="L2"	6933. .6994			
repeat_region	/rpt_family="ERV1"	6995. .7535			
repeat_region	/rpt_family="ERV1"	7516. .7619			
repeat_region	/rpt_family="ERV1"	7620. .7921			
repeat_region	/rpt_family="Alu"	7922. .8073			
repeat_region	/rpt_family="ERV1"	8078. .8386			
repeat_region	/rpt_family="Alu"	8456. .8893			
repeat_region	/rpt_family="ERV1"	9675. .9781			
repeat_region	/rpt_family="MERL"	10536. .10774			
repeat_region	/rpt_family="MERL"	10784. .11233			
repeat_region	/rpt_family="ERV1"	11428. .11544			
repeat_region	/rpt_family="MERL"	11545. .11899			
repeat_region	/rpt_family="MERL"	11900. .11920			
repeat_region	/rpt_family="MERL"	12078. .12359			
repeat_region	/rpt_family="Alu"	12420. .12706			
repeat_region	/rpt_family="MERL"	14365. .14403			
repeat_region	/rpt_family="AT-rich"	14599. .14928			
repeat_region	/rpt_family="MERL_type"	14930. .15505			
repeat_region	/rpt_family="ERV1"	15849. .16157			
repeat_region	/rpt_family="Alu"	17559. .17652			
repeat_region	/rpt_family="MIR"	18143. .18624			
repeat_region	/rpt_family="ERV1"	18625. .19034			
repeat_region	/rpt_family="ERV1"				

US-09-856-836-2 (1-351) x AC010878 (1-154115)					
QY 1 MetalMetarglInthrProleuthrCysSerglyHisThrArgProvalasp*** 20					
Db 94882 ANGGCAATGACACAGCGCTGCTACCTTTGGCCACACGTGGCCATGGTGTGTTGG 948233					
QY 21 AlapheserglytlethrprotyrgrlytyrphellelleserAlacyslaspgllylys 40					
Db 94822 GCCCTGAGTGCATCCAGCACTTATGGGTATTTCTTAACCGCGCTTACAAAGTGTAA 947633					
QY 41 PrometleuargglngllyasprthrgrlyasprtrpilleglythrPhelenglyhislysgly 60					
Db 94762 CCTATGCTACGCCAGGGAATATCAGAGACTGGATTGAAACCTTTGGGCTATTAAGGT 947033					
QY 61 AlavaltrpGlyAlaThrLeuasnlyAspAlaThrLysAlaAlaThrAlaAlaasp 80					
Db 94702 GCGTTGGGGTGCACACTGAT-----GCCACCAAGCAGCAACGACCTGCAGAT 94649					
QY 81 PheethAlaLysValTtrpAspAlaValserGlyAspGluLeuMetThrLeuAlaHislys 100					
Db 94648 TTCACAGCCAAAGTGGGATCTGCTTACAGAGATGAATGATGACCTTGGCTATAAA 94589					
QY 101 HisileVallysthrValaspphrthrgrlnaspsersAspTyleu***thrglyglyln 120					
Db 94588 TACATTGTCAAGACTGTGATTTTCATGACGAGTATGATTAATTAATTTGTTAACCAAGGACAG 94529					
QY 121 AsplylsleuLeuArgtlleTyraSpleuasnlysProgluAlaGluProlysgluLieser 140					
Db 94528 GATTAACCTTTCAGCATGTATGACTTGAAACCAACCAAGAACCTTACGAAATTAAGT 94469					
QY 141 GlyHisThrSerglylylelylys---AlaLeuTrpCysSerAspAspLysGlnleu 159					
Db 94468 GGTACACACTTCGTATATAAAAAAGCGCTCATGTGTGACAGTGAAGATTAACAGATTCCT 94409					
QY 160 SerAlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSer 179					
Db 94408 TCTGCTGATGACAAACTGCTGACTTTGGATCAAGTACTATGACAGAAAGTGAATCT 94349					
QY 180 LeuAsnPhaasmetSerValserSerMetGluTyrlleProgluGlyGluLeuVal 199					


```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0447c21
----- Summary Statistics -----
Sequencing vector: M13; 63%
Sequencing vector: plasmid; 37%
Chemistry: Dye-terminator; 48% of reads
Chemistry: Dye-primer; 52% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 181938 bases at least Q40
Consensus quality: 183284 bases at least Q30
Consensus quality: 184219 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 186317; sum-of-contigs
Quality coverage: 6.62 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*
*
1      4128: contig of 4028 bp in length
*      4029      4028: contig of unknown length
*      4129      7967: contig of 3839 bp in length
*      7968      8067: gap of unknown length
*      8068      16351: contig of 8284 bp in length
*      16352      16451: gap of unknown length
*      16452      26437: contig of 9986 bp in length
*      26438      26537: gap of unknown length
*      26538      35887: contig of 9351 bp in length
*      35889      35888: gap of unknown length

```

FEATURES	*	35989	56349: contig of 20361 bp in length
	*	56350	56449: gap of unknown length
	*	56450	85530: contig of 29081 bp in length
	*	85531	85630: gap of unknown length
	*	85631	187017: contig of 101387 bp in length.
source			Location/Qualifiers
		1..187017	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="7"	
		/clone="RP11-447C21"	
		1..4028	
misc_feature	/note="assembly_name:Contig26"		
	4129..7967		
misc_feature	/note="assembly_name:Contig27"		
	8068..16351		
misc_feature	/note="assembly_name:Contig28"		
	16452..26437		
	/note="assembly_name:Contig29		
	clone_end:77		
	vector_side:left"		
misc_feature	26538..35888		
	/note="assembly_name:Contig30"		
misc_feature	35989..56349		
	/note="assembly_name:Contig31"		
	56450..85530		
misc_feature	/note="assembly_name:Contig32		
	clone_end:SP6		
	vector_side:right"		
misc_feature	85631..187017		
	/note="assembly_name:Contig33"		
BASE COUNT	53793 a 37172 c 37140 g 58180 t	732	others
ORIGIN			
Alignment Scores:			
Pred. No.:	4.84e-110	length:	187017
Score:	1334.50	Matches:	291
Percent_Similarity:	66.81%	Conservative:	13
Best Local Similarity:	63.96%	Mismatches:	44
Query Match:	72.53%	Indels:	110
DB:	2	Gaps:	3
US-09-856-836-2 (1-351) x AC010101 (1-187017)			
QY	1	MetaMeta.rglinthProleuthrcysserglyhishrharProvalaIasp***	20
Db 100587	AWAGCAATGAACACGATGCTCTCACCGTTCGGCCACATATGCGCCGTCGACTTG		100528
QY	21	AlaPheSerGlyIlethPro-	27
Db 100527	GCCTTCAGTGGCAATCATGCTTTAGGGGCTTTTTCCTCCACAGACAGAAATCTTGCTGTG		100465
QY	27		27
Db 100467	TCGCCAGGCTGAGTGCATGTGTGCATCTTGCTCAGTCAAACTCCGGCTCCAGGT		100408
QY	27		27
Db 100407	TCAGTGAATTCTCTGCTCCCTCCAGCTCCCGAGTAGCTGGATTACAGCGACGCTATCAC		100348
QY	27		27
Db 100347	TCCTGGCTAATTTTGTATTTTATAGTAGAGACGGGGTTTCACCATGTTGGCCCAAGCTGTT		100288
QY	27		27
Db 100287	CTGCAATCTTGACCTCAGGTGATCCGCTGCTTGCCCTCCAAAGTCTGGGATTACA		100228
QY	28	-----TyrGlyTyrPheLeuIleSerAlaCysIysAs	38
Db 100227	GGCGTGAAGCCACTCGCGCTGGCCACGCTTATATGAGATATTTCTAAATCAGTGGCTTCAAAAG		100168
QY	38	pGlyIysPheLeuArgIleGlnGlyAspThrGlyAspTyrIleGlyTyrPheLeuGlnYH	58

```

|||||
Db 100167 CGGCAACCTATGCTATGCGGAGATACAGAGACTGATGGACAAATTTGGCTCA 100108
|||
Qy 58 slysglyalvaltprgylalathrleuanslyspalathrlysalalathr1a1 78
|||
Db 100107 TAAAGATGCTGTCGAATGCAACACTAAATAGATGCGCACCAAGCGACAGACAG 100048
|||
Qy 78 aalaspheprralalyvalttrpasplalalserglyaspglulemetthrleu1 98
|||
Db 100047 TGCAGATTTGACA---AGTG-TGGATGCTATGCCAGAGATGAACTGACCTGGC 99992
|||
Qy 98 ahlsyshtsilevallysthrvalaspheprrhlnaspserasnryleu***thr1 118
|||
Db 99991 TCATTAACATTAATCAACACATGATTTCAATGAGATAGTAATATTTGATACCA 99932
|||
Qy 118 yglylnasplysleuauarglletyrasplesuanslysproglyualgluprolys1 138
|||
Db 99931 GGCAGAGATTAACCTGTTTGCATATATGACTTGAACAAACCCGAGACAACTGAG 99872
|||
Qy 138 ulleserghlysthrserglylylelysalaleutrpysseraspplsglnt1 158
|||
Db 99871 AATTGATGATCAGCTGATTAATAAGCCCTATGCTACAGATAGATTAACAGAT 99812
|||
Qy 158 eluseralaspasplysthrval-Argleutrpasphsalathrmetthr1gluval1 178
|||
Db 99811 TCTTCTCTGATGACAAACAGTTAAGACTTGGATCATGCTACTATGACAGAGTGA 99752
|||
Qy 178 yserleuanspshasmetservalsermetlutyrlleprogluglylul1el 198
|||
Db 99751 AATCTCTAAATTTTATATGATGCTTCCAGTAGTAAATATATTTCTGAGGAGATTT 99692
|||
Qy 198 euvallethrlyrlygylasergserlialaphehiserlialserleugluprol1el 218
|||
Db 99691 TGTATTAATCTATGCTGCTGATCCAGTCTTTACAGTACAGTATTTGAGAACATTA 99632
|||
Qy 218 yserpheglualaproalathr1leasnseralaser***HisProglulysgluphel 238
|||
Db 99631 AATCTTCAAGCTCCGCAACATTAACCTGATCTTCCACTTCCAGTCAAGAAATTTTC 99572
|||
Qy 238 euvalnalaglygluaspphlyleutlyrlysthrlyrlyasnserglylul1glul 258
|||
Db 99571 TTTGTTCCAGTGTGAAGATGTTAACT-TATAGATGATGATTAATAGTGGAGAAAT 99513
|||
Qy 258 engluserlyrlysglyhlyspheglproliethiscysvalatrphepserproasp1yg 278
|||
Db 99512 TAGAATCCATCAAGAGACTGCTGCTCTATTCAGTGTGAGATTTGATCCGATGAG 99453
|||
Qy 278 lulautryalaserglysergluasplythrlleuargleutrpclnthval1gl1yl 298
|||
Db 99452 AACTCTATGCGCAGCTGCTGGAAGATGGAACACTGAGAC--TGGCAAACTGTGGTAGAA 99395
|||
Qy 298 yathrlyrlyleutrp1yrcysval***Proglugluaspserglyluleu1alasp 318
|||
Db 99394 AAACATATGCTCTTTGGAATGTGCTTCTTCAAGAGATGATGCTGAGCTGCAAGC 99335
|||
Qy 318 rolyslleeglyphepogluthr1alaglulglul1glul1al1alaserglua 338
|||
Db 99334 CAAGAATTTGTTTCCAAAGACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 99281
|||
Qy 338 snseraspserlletyrsersthrpogluvallysal1 351
|||
Db 99280 ATTTCAGATTCATCTATCTTCAACTCTCAAGTTAAGGCC 99240
|||
RESULT 11
AC010132/ 166680 bp DNA linear PRI 30-SEP-2000
DEFINITION Homo sapiens BAC clone RP11-111K18 from 7p11.2-p2, complete
sequence.
ACCESSION AC010132
VERSION AC010132.5 GI:10440742
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

```

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
PUBLISHED
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 166680)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 166680)
Andrews, S., Wohlmann, P. and Scherzger, E.
The sequence of Homo sapiens BAC clone RP11-111K18
Unpublished
3 (bases 1 to 166680)
Waterston, R.H.
Direct Submission
Submitted (13-SEP-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 166680)
Waterston, R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 30, 2000 this sequence version replaced gi:8954217.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplen@wustl.edu
----- Summary Statistics
Center project name: H_NH0111K18

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/GRF/CH7>, send
<mailto:greene@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frenken, E.,
Tateno, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is GS1-308H5, 200 bp overlap.
Actual start of this clone is at base position 156485 of GS1-308H5;
actual end is at base position 166680 of RP11-111K18.

There are polymorphic base pair differences in the overlap between
the clone RP11-111K18 and GS1-308H5.

FEATURES
SOURCE

```
Location/Qualifiers
1..166680
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7p11.2-p2"
/clone="RP11-111K18"
/clone_lib="RP11-11"
2..397
/rpt_family="ERV1"
repeat_region
1478..1785
/rpt_family="Alu"
1807..1913
/rpt_family="L1"
2138..2508
/rpt_family="ERV1"
2532..2817
/rpt_family="MALR"
2846..2911
/rpt_family="ERV1"
2912..3043
/rpt_family="Alu"
3044..3115
/rpt_family="ERV1"
3087..3598
/rpt_family="ERV1"
3600..3780
/rpt_family="ERV1"
3781..4090
/rpt_family="Alu"
4091..4257
/rpt_family="ERV1"
4260..4865
/rpt_family="ERV1"
4902..4994
/rpt_family="ERV1"
6077..6610
/rpt_family="ERV1"
7185..7327
/rpt_family="ERV1"
7327..7539
/rpt_family="ERV1"
7579..8074
/rpt_family="ERV1"
8169..9081
/rpt_family="ERV1"
9082..9362
/rpt_family="Alu"
9363..9633
/rpt_family="ERV1"
9634..9943
/rpt_family="Alu"
9944..10150
/rpt_family="ERV1"
10151..10567
/rpt_family="MALR"
10568..12227
/rpt_family="ERV1"
12316..12514
/rpt_family="ERV1"
12532..12582
/rpt_family="ERV1"
13068..13614
/rpt_family="ERV1"
15167..15445
/rpt_family="Alu"
16828..17212
/rpt_family="MALR"
17548..17857
/rpt_family="Alu"
17897..18030
/rpt_family="ERV1"
18365..18652
```

```
repeat_region /rpt_family="Alu"
19118..19746 /rpt_family="ERV1"
repeat_region /rpt_family="ERV1"
19750..19814 /rpt_family="ERV1"
repeat_region /rpt_family="ERV1"
19852..20169 /rpt_family="ERV1"
repeat_region /rpt_family="ERV1"
20224..20412 /rpt_family="ERV1"
repeat_region /rpt_family="ERV1"
20540..20840 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
20879..20997 /rpt_family="MIR"
repeat_region /rpt_family="MIR"
21649..21848 /rpt_family="MERL_type"
repeat_region /rpt_family="MERL_type"
22158..22177 /rpt_family="MIR"
repeat_region /rpt_family="MIR"
22178..22488 /rpt_family="ERV1"
repeat_region /rpt_family="ERV1"
22489..22777 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
22778..22919 /rpt_family="ERV1"
repeat_region /rpt_family="ERV1"
22920..22979 /rpt_family="MIR"
repeat_region /rpt_family="MIR"
23031..23067 /rpt_family="L2"
repeat_region /rpt_family="L2"
23518..23928 /rpt_family="MALR"
repeat_region /rpt_family="MALR"
24091..25047 /rpt_family="L2"
repeat_region /rpt_family="L2"
25094..25312 /rpt_family="L2"
repeat_region /rpt_family="L2"
25375..25549 /rpt_family="L2"
repeat_region /rpt_family="L2"
25663..25817 /rpt_family="MERL_type"
repeat_region /rpt_family="MERL_type"
27098..27212 /rpt_family="L2"
repeat_region /rpt_family="L2"
27657..27925 /rpt_family="MALR"
repeat_region /rpt_family="MALR"
28020..28061 /rpt_family="L2"
repeat_region /rpt_family="L2"
28527..28780 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
29058..29361
```

Alignment Scores:

```
Pred. No.: 1,49e-109 Length: 166680
Score: 1328.50 Matches: 290
Percent Similarity: 66.59% Conservative: 13
Best Local Similarity: 63.74% Mismatches: 45
Query Match: 72.20% Indels: 110
DB: 9 Gaps: 3
```

US-09-856-836-2 (1-351) x AC010132 (1-166680)

```
OY 1 MetAlaMetArgGlnThrProLeuThrcysSerGlyHisThrArgProValAlaAsp*** 20
Db 48948 ATAGCAATGAGACTGATGCTGCTACCTGCTGCGACATATGATGCGCTGACTTG 48889
OY 21 AlaPheserGlyIleThrPro----- 27
Db 48888 GCCTTCACTGTCATCATGCTTGAAGGGCTTTTTCCTCCAGACAGACATCTGCTCG 48829
OY 27 ----- 27
Db 48828 TCGCCAGCGCTGAGTGCAGTGTGCATCTGCTCACTGCAAACTCCGCTCCAGGT 48769
OY 27 ----- 27
Db 48768 TCAAGTATTTCTCTGCTGCTCAGCTCCCGAGTATGCTGGATTACAGGCGCGCTATTCAC 48709
```

```

OY 27 ----- 27
Db 48708 TCCTGCTAATTTTGTATTTTGTAGAGACGGGTTTCACCATGTGGCCAGCTGT 48649
OY 27 ----- 27
Db 48648 CTCGAACCTTCGACCTGAGGTGATCCGCTGCTGCCCTCCCAAAAGCTGGGATTTACA 48589
OY 28 ----- TyrglytyrPheLeuIleSerAlaCysIysAs 38
Db 48588 GGGGTGAGGCACTGGGCTGGCCAGCTTATAGTATTTCTTAATCATGCTCTGCAAGA 48529
OY 38 pGlyIysPromeLeuArgInglIysPthrGlyAspTrpIleGlyThrPheLeuGlyH 58
Db 48528 CGGCAAACTATGCTATGCGAGGAGATACAGAGACTGATTTGGAACAATTTTGGGTCA 48469
OY 58 slysgIyAlaValTrpGlyValaThrLeuAsnIysAspAlaThrIysAlaAlaThrAlaAl 78
Db 48468 TAAAGATGCTGTGGAATGCAACACTGAAATTAAGATGCCACCAAGACACCCACAGCAGC 48409
OY 78 aAlaAspPheThrAlaIysValTrpAspAlaValSerGlyAspGluIleuMetThrIleuAl 98
Db 48408 TCGAGATTTTCACA---AAGTG-TGGATGCTATCCAGAGATGAACTGATGACCTGGC 48353
OY 98 aHisIysHisIleValIlysthrValAspPheThrGlnAspSerAsnTrpLeu**ThrG 118
Db 48352 TCATAAACACATATATCAGACTATGATTTTCATGACAGATAGTAATTTATTGTTAACAG 48293
OY 118 yGlyGlnAspIysLeuLeuArgIleTyrrAspLeuAsnIysProGluAlaIubProIysG 138
Db 48292 GGGACGAATTAACCTTTTGGCATATGACTGAAACAAACCCGAAAGCACACACCTGGGA 48233
OY 138 uIleSerGlyHisThrSerGlyIleIysIysAlaLeuTrpCysSerAspIysGlnI 158
Db 48232 AATTAGTATGATCAGCTGATATATAAAGGCCCTATGGTACAGAGATTAACAGAT 48173
OY 158 eIeuSerAlaAspAspIysThrVal-ArgIeuTrpAspHisAlaTrpMetThrGluVal 178
Db 48172 TCTTTTCTTGATGACAAACCTGTAAAGCTTTGGATCTATGCTATGACGAAGTGA 48113
OY 178 ySerLeuAsnPheAsnMetSerValSerSerMetGluTrpIleProGluGluIle 198
Db 48112 AATCTTAATATTTTATATGTTGGCAGTAGTAGTATGAAATTTTTCGAGGAGAGATTT 48053
OY 198 euValIleThrTrpGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIle 218
Db 48052 TGGTAAATTAATTAATGCTGATCCAGTCTTTTACAGTACAGTAAGTTTGAACCAATTA 47993
OY 218 ySerPheGluAlaIleProAlaThrIleAsnSerAlaSer**HisProGluIysGluPhe 238
Db 47992 AATCTTTCAAGCTCTGCAACATTAATCTTAATCTTTCACCTTGAGAAAGAAATTTTC 47933
OY 238 euValAlaGlyGlyGluAspPheIysLeuTrpIysTrpAspTrpAsnSerGlyGluIle 258
Db 47932 TTGTTTCAGGTGTGGAAGATGTTAAACT-TATTAAGTATGATGATTAAGTGAAGAAT 47874
OY 258 euGluSerTrpIysGlyHisPheGlyProIleHisCysValaIrgPheSerProAspGly 278
Db 47873 TAGAATCTTACAAAGACACTGCTGCTATTCATCTGAGATTTGATCTCATGATGAG 47814
OY 278 IueuTrpAlaSerGlySerGluAspGlyThrIleuArgLeuTrpGlnThrValValGly 298
Db 47813 AACTTATGCAAGTGTCTGGAAGATGGAACACTGAGAC--TGGCAAACTGTGGTGAAG 47756
OY 298 ySthTrpGlyLeuTrpIysCysVal**ProGluGluAspSerGlyGluIleAlaIalysp 318
Db 47755 AATCATATGCGCTTTGGAATATGCTGCTTCTGAAAGATTAAGTGTGACTGCGCAAGC 47696
OY 318 rOlIysIleGlyPheProGluThrAlaGluGluIleuAlaGluGluIleAlaSerGlu 338
Db 47695 CAAGAGTGTGTTTCCAAAGACAAAGAGAGAG--CTAGAAGAAATTTGCTCAGAGA 47642
OY 338 snSerAspSerIleTyrrSerSerThrProGluValIysAla 351

```

```

Db 47641 ATTGAGATTCACATCTTCTTCAACTCTCCAGTTAAGGCC 47601
RESULT 12
AC020884_3/c
WPCOMMENT
Sequence split into 4 fragments
Fragment Name Begin End
AC020884_0 1 110000
AC020884_1 100001 210000
AC020884_2 200001 310000
AC020884_3 300001 373038
Continuation (4 of 4) of AC020884 from base 300001 (AC020884 Mus musculus clone RP23-
Alignment Scores:
Pred. No.: 1, 12e-91 Length: 73038
Score: 1128.00 Matches: 263
Percent Similarity: 80.29% Conservative: 14
Best Local Similarity: 76.23% Mismatches: 52
Query Match: 61.30% Indels: 22
DB: 2 Gaps: 5
US-09-856-836-2 (1-351) x AC020884_3 (1-73038)
OY 1 MetalameTarGlnThrProLeuThrCysSerGlyHisThrArgProValIalAsp** 20
Db 25489 GTGGCCATGAGGAGACTCTGCTACATTCCTCAGGCCGCGCATGCGGCTGTGGATTTG 25430
OY 21 AlaPheSerGlyIleThrProGlyTrpPheLeuIleSerAlaCysIysAspGlyIys 40
Db 25429 TCCCTCAGCAGCATCCCGCTAAAGCTGCTTCTGATCAGTCTGCAAGATGGCCAA 25370
OY 41 PromeLeuArgInglIysPthrGlyAspTrpIleGlyThr-PheLeuGlyHisIysG 60
Db 25369 CTAATGCTCACACAGGAGATACAGAGACTAGATTGGAACATTTTGGGTATAAAG 25310
OY 60 yAlaValTrpGlyAlaThrIleu---AsnIysAspAlaThrIysAlaAlaThrAlaAl 79
Db 25309 TGGTGTGGGGTACAACATTTGACTTAATAAGATGCCCAAAACCTGTACACAGCTGC 25250
OY 79 aAspPheThrAlaIysValTrpAspAlaValSerGlyAspGluIleuMetThrIleuAl 99
Db 25249 AGACATTCACA-----GTGATGATGCTGCTCAGGAGATGAAATGATACCCGCTCA 25196
OY 99 slyshsIleValIysThrValAspPheThrGlnAspSerAsnTrpLeu**ThrGly 119
Db 25195 TCAAGCAATTTGCAAGACTGTGAAGTTCACACAGAGATGAGATTAACCTGCTGAGCA 25136
OY 119 yGlnAspIysLeuLeuArgIleTyrrAspLeuAsnIysProGluAlaIubProIysG 139
Db 25135 -CAGGATTAAGTGTGACATACAGACTGGA-AAACCTGAAAGCACACCTTAAGGAAT 25078
OY 139 eSerGlyHisThrSerGlyIleIysIysAlaLeuTrpCysSerAspIysGlnIle 159
Db 25077 TAATGCTCACACTTCGGTATTAAGAGCTCTGTGTCAGATGAATATAATAGTCTCT 25018
OY 159 uSerAlaAspAspIysThrValArgIleuTrpAspHisAlaThrMetThrGluValIys 179
Db 25017 TTCAAGTATGATTAATAAAGCTTTGGCTGCTGCAATACACAGAGTACAGAAAGTAATC 24958
OY 179 rIleuAsPheAsnMetSerValSerSerMetGluTrpIleProGluGluIleuVal 199
Db 24957 TTAATAATGT-AATATGCTTTT-AGCAGACACAAATATATATTCCTGAGGAGATCTTGGT 24900
OY 199 lIleThrTrpGlyArgSerIleAlaPheHisSerAlaValSerLeu-----GluPr 216
Db 24899 T-----ATGCTTTTCAATGTCAGTAATATTAATTTATGTCGGAGGCC 24858
OY 216 oIleIysSerPheGluAlaIleProAlaThrIleAsnSerAlaSer**HisProGluIys 236
Db 24857 AATTAATCTTGTGAAGCTCTGCGACCATCAATTTGTCATCTCTCATCAGAGAGAGA 24798
OY 236 u-PheLeuValAlaGlyGlyGluAspPheIysLeuTrpIysTrpAspTrpAsnSerGly 256

```



```

FEATURES
SOURCE
    1. location/Qualifiers
    1..226520
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /chromosome="11"
    /clone="RP23-128D9"
    /clone_1lb="RPC1-23"
BASE COUNT    64093 a 48521 c 50155 g 63751 t
ORIGIN

```

Alignment Scores:	Pred. No.:	4,29e-91	Length:	226520
Score:	1128.00	Matches:	263	
Percent Similarity:	80.29%	Conservative:	14	
Best Local Similarity:	76.23%	Mismatches:	52	
Query Match:	61.30%	Indels:	22	
DB:	10	Gaps:	5	
US-09-856-836-2 (1-351) x AL645802 (1-226520)				
QY	1	MetaLametaRglnhtrProLeuThrCySerGlyHisThrArgProValIValAsp***	20	
Db	142443	GTGGCCATGAGGACAGACTCGTCACCTTGCTCAGGCCGCCATCCGGCTGTGTGGATTTC	142384	
QY	21	AlaPheSerGlyLeuThrProTyrGlyTyrPheLeuIleSerAlaCysAlaSpglyLys	40	
Db	142383	TCCTTGACGAGCATCCCGCTTAAAGCGTCCTTTCGATCAGTGTGCGCAAGAATGGCCAA	142322	
QY	41	ProMetLeuArgGlnGlyAspThrGlyAspTrpIleGlyThr-PheLeuGlyHisLysG1	60	
Db	142323	CCATGCTCCACAGGAGAAATACAGAGACTGATGATGGAAATTTTGGGTCTAAAGG	142264	
QY	60	YAlaValITrpgIyAlaThrLeu--AsnLysAspAlaThrLysAlaIaThrAlaIaAl	79	
Db	142263	TGCTGTTTGGGGACAACTTACTAATAAAGATGCCACCAAACTGCTACAGCAGCTGC	142204	
QY	79	AspPheThrAlaLysValITrphAspAlaValSerGlyAspGluLeuMetThrLeuAlaH1	99	
Db	142203	AGCAATCAACA-----GCTGTGATGAGCTGTCTCAGAGATGAAATGATGACCCGTCTCA	142150	
QY	99	SlyHisIleValLysThrValAspPheThrGlyAspSerAsnTyrLeu**ThrglyG1	119	
Db	142149	TCAGCAGATTGTCAAGACTGTGAAGTTACACAGAGATAGCAGTTACTGCTGACTGGCA	142090	
QY	119	YglAspLysLeuLeuAlaGlyIleTyrAspLeuAsnLysProGluAlaGluProLysGluI1	139	
Db	142089	-CAGGATAAACGTTGGACATACACGACTTGA--AAACCTGAAGCAGAACTTAAGGAAT	142032	
QY	139	eSerGlyHisThrSerGlyIleLysValAlaLeuTrpCysSerAspAspLysGlnIleLe	159	
Db	142031	TATGCTGCACACTTCGGTATTAAAAGGCTGTGTGTGCAAGTGAATAAATGTTCT	141972	
QY	159	uSerAlaAspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSe	179	
Db	141971	TTTCACGTATGATTAACACTGTTGGCTGTGGCATACACAGTACAGTAAGTGAATC	141912	
QY	179	TrpLeuSnpHeaSnkSetSerValSerSerMetGluTyrIleProGluGlyIleLeuVa	199	
Db	141911	TTTAAATGT-AATATGCTTT-ACCACACCAAAATATATATTCGAGGAGAGATCTGGGT	141854	
QY	199	IleThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeu-----GluPr	216	
Db	141853	T-----ATTGCTTTCAATGTGCGAGTATTAATTAATTTGCCGAGGC	141812	
QY	216	oileLysSerHeGluAlaProAlaThrIleAsnSerAlaSer**HisProGluLysG1	236	
Db	141811	AATTAAATCTTTGAACCTCTGGACCATCAATCTTGATCTCTTCATCCAGAGAAGA	141752	
QY	236	u-PheLeuValAlaGlyG1GluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyG	256	
Db	141751	ACTTTTGTGTTCGAGGTGGAGAAC--TTTAACTGTATTAAGTCTTAAATGGGTGGAG	141693	
QY	256	IuGluLeuGluSerTyrLysGlyHisPheGlyProIleAsCysValArgPheSerProA	276	
Db	141692	AAGAAATTAGATCTTACAAAGGTCAATGTCATCAGTGTG--AGATTCACTCTG	141635	
QY	276	spGlyGluLeuTyrAlaSerGlySerGlyHisAspGlyThrLeuArgLeuTrpGlnThrVal	296	
Db	141634	ATGTGTAACCTATACCACTGGCTCTGAAGAATGGGACACAGGAATTTGGCCAAACTGTGG	141575	
QY	296	aleGlyLysThrTyrGlyLeuTrpLysCysVal***ProGluLysAspSerGlyGluIleVa	316	

Db 101920 gcc 101922

RESULT 15
AY061000
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY061000
Drosophila melanogaster
AY061000
AY061000.1
GI:16768657
FLI-CDNA.
Drosophila melanogaster.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1591)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,
Gonzalez, M., Gartin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,
Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,
Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.
Direct Submission
Submitted (30-OCT-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

TITLE
JOURNAL

COMMENT

This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy. Presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unsplined precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES
source
1..1591
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="59E3-59E3"
1..1591
/gene="CG3957"
/note="alignment with genomic scaffold AB003461"
/db_xref="FLYBASE:FBgn0034876"
345..1331
/gene="CG3957"
/note="longest ORF"
/codon_start=1
/product="HL01517p"
/protein_id="AA128548.1"
/db_xref="GI:16768658"
/db_xref="FLYBASE:FBgn0034876"
/translation="MAANNLRQIPITCSGHTRPVHLDSPICDAGYFLISACKDGS
MLRGDGDWGVTEGEGHGAVMATLNRNATLASAGDFGKVMNVTGAEHISPOH
KHIVSAFDRDSENITVGSNEKLVRFVLEOPAEDEYAGHGAIKRALFCGDK
IIISADEKTVLRMTGIEYORLOFNSNPSELTSSDNHTITSHSSSTISFWIDTL
KKLEKVPPTNVAASLHPDKHYVCGGDFKMTKTYITGNETESKGRHGPVHSVK
FSPDGLYASSESGITRLMOTIVGKTYGTMKCTEPADLNSISSPREVAVN"

BASE COUNT
ORIGIN
436 a 451 c 425 g 279 t

Alignment Scores:
Pred. No.: 9.18e-87 Length: 1591

Score: 1052.50
Percent Similarity: 69.07%
Best Local Similarity: 53.87%
Query Match: 57.20%
DB: 3
Gaps: 3

US-09-856-836-2 (1-351) x AY061000 (1-1591)

QY 3 MetArgInThrProLeuThrCysSerGlyHisThrArgProValAlaLysPhe***AlaPhe 22
DB 360 TTGGAGCAAAATCCCGTGGACCTTCAGTGGCCACACCCGCTGTCACCTGACCTTC 419
QY 23 SerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysPheGlyLysProMet 42
DB 420 AGCGACATCTGCGACCCCGCTTCTTCATCTCGGCTGCAAGATGCGACCGCATG 479
QY 43 LeuArgGlnGlyAspThrGlyAspTyrIleGlyThrPheLeuGlyHisLysGlyAlaVal 62
DB 480 CTCGGCAGCGCGACACCGCGCTGTTGGGACCTTCGAGGAGACACAGGCGCGCTG 539
QY 63 TTPGLYAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThr 82
DB 540 TGGACGCCACGCTGACCGCAATGCGACCTGCGCGCTCGGAGCGCGCTTCAC 599
QY 83 AlaLysValTyrAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLysHisIle 102
DB 600 GCGACAGTGTGGATGCGGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 659
QY 103 ValLysThrValAlaAspPheThrGlnAspSerAsnTyrLeu***ThrGlyGlyLysPhe 122
DB 660 GTGAAAGGCGTGGCTTCACGAGGATTCGAGACATAGTACTGCGAGCAAGCAAG 719
QY 123 LeuLeuArgIleTyrAspLeuAsnLysProGluIleGluIleGluIleSerGlyHis 142
DB 720 CTGTCGCGGCTTTCACCTGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 779
QY 143 ThrSerGlyIleLysAlaLeuTyrCysSerAspLysGluIleLeuSer***Ala 161
DB 780 AGCGGGCGCATCAACCGCGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 839
QY 162 AspAspLysThrValArgLeuTyrAspHisAlaThrMetThrGluValLysSerLeuAsn 181
DB 840 GAGGACAGACGCTGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
QY 182 PheAsnMetSerValSerSerMetGluTyrIleProGluGluIleLeuValIleThr 201
DB 900 TTCAACAGCAATCCGACAGCGCTGAGATCTCCAGCGCAACACACATRACATACG 959
QY 202 TyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPheGlu 221
DB 960 CACGGCTCTCGATTAAGCTTCTGGAGATCGACACGCGGAGAGAGCGGAGAGAG 1019
QY 222 AlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAlaGly 241
DB 1020 GTGCGAGCAAGACGTGGCGTGGCGGATTCGATCCGGATTAAGCATGTGTTGTTGGC 1079
QY 242 GilyLysAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluIleGluLysTyr 261
DB 1080 GCGCGAGCACTTAAGATGTACAAATTCGACATCAACAGGAAGCAAAATTAATCTTTC 1139
QY 262 LysGlyHisPheGlyProIleHisCysValAlaArgPheSerProAspGlyLysLeuTyrAla 281
DB 1140 AAGGAGCACTTTCG 1199
QY 282 SerGlySerGlyAspGlyThrLeuArgLeuTyrProGluThrValAlaGlyLysThrTyrGly 301
DB 1200 AGCGGCTCCGAGGAGCGGACATCTCGGCTGTGGAGACACACCGTGGCAAGACCTAGG 1259
QY 302 LeuTyrLysCysVal***ProGluGluAspSerGlyGluLeuAlaLysProLys----- 319
DB 1260 TTGTGAAGTGCACGAGCGCGCGGAGCTGCGCAATTCATCAGCTCGCGCGCGAGG 1319
QY 319 ----- 319

Mon Feb 10 11:12:50 2003

us-09-856-836-2.rge

Page 18

Dd	1320	CAGGCAAACTGATGTTCGGCCACAGGAATCGTGTTACACAGCCACAGCTGCACC	GCG	1379
Oy	320	-----TlEgLyPheProGLuThrAlaLglLglLevalAglUglU---		333
		::: :::		
Dd	1380	ACGCAACAAGGAGCATTGATTCGGAGCTACGGCTGCAATCCGATTTTCACACACAGA		1439
Oy	334	-----lleAlSerGlusAnserIleTyrrSerSer		345
		::: :::		
Dd	1440	CAACTCGTAGTCGGCGCTAC-AATAAGCATTTCTCATATTCCAAC		1483

Search completed: February 6, 2003, 23:30:17
Job time : 3659 secs

•
•
•
•

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2003, 13:06:48 ; Search time 2160 Seconds

(without alignments)
2631.767 Million cell updates/sec

Title: US-09-856-836-2

Perfect score: 1840

Sequence: 1 MAHQPTLTCGHTRPVVDX.....EELASSENSDIYSTPEVKA 351

Scoring table:

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q/cgn2_1/USFTO.spool/US09856836/runat_31012003_144238_22208/app_query.fasta_1.519

-DB=EST -QPM=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USRR=US09856836.ecgn.1.1.763.ernat.31012003_144238_22208 -NCPU=6 -ICPU=3

-NO_XLRY -NO_MMP -IARECOVER -NRG_SCORES=0 -WAIT -LONGLOG -DEV_TTIMEOUT=120

-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YCAPOP=10 -YCAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST :

1: em_estbda:*

2: em_estbda:*

3: em_estbda:*

4: em_estbda:*

5: em_estbda:*

6: em_estbda:*

7: em_estbda:*

8: em_estbda:*

9: em_estbda:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1784.5	97.0	1860	11	AF161496	AF161496 Homo sapi
2	1422	77.3	904	11	AU123736	AU123736 AGENCOURT
3	1380.5	75.0	1174	14	BM802585	BM802585 AGENCOURT
4	1325	72.0	944	9	AL517952	AL517952 AGENCOURT
5	1313.5	71.4	1055	13	BM472016	BM472016 AGENCOURT
6	1302.5	70.8	880	14	BQ219033	BQ219033 AGENCOURT
7	1267	68.9	902	14	BQ889506	BQ889506 AGENCOURT
8	1265	68.8	875	14	BQ439153	BQ439153 AGENCOURT
9	1235	67.1	757	12	BG687473	BG687473 AGENCOURT
10	1228.5	66.8	822	9	AJ395256	AJ395256 AGENCOURT
11	1217.5	66.2	1019	13	BM450309	BM450309 AGENCOURT
12	1203	65.4	800	9	AU123939	AU123939 AGENCOURT
13	1193	64.8	796	12	BG778828	BG778828 AGENCOURT
14	1184	64.3	901	9	AL556515	AL556515 AGENCOURT
15	1182	64.2	1086	14	BQ433088	BQ433088 AGENCOURT
16	1179	64.1	901	9	AL518570	AL518570 AGENCOURT
17	1176.5	63.9	879	14	BQ222878	BQ222878 AGENCOURT
18	1173	63.7	881	9	AU132581	AU132581 AGENCOURT
19	1150	62.5	1038	9	AL545602	AL545602 AGENCOURT
20	1141.5	62.0	704	12	BG740119	BG740119 AGENCOURT
21	1139	61.9	885	14	BQ883467	BQ883467 AGENCOURT
22	1130.5	61.4	874	9	AL570903	AL570903 AGENCOURT
23	1121	60.9	748	13	BI090693	BI090693 AGENCOURT
24	1115	60.6	951	9	AL551950	AL551950 AGENCOURT
25	1115	60.6	959	9	AL547128	AL547128 AGENCOURT
26	1113.5	60.5	833	9	AU119614	AU119614 AGENCOURT
27	1108	60.2	962	9	AL539156	AL539156 AGENCOURT
28	1106	60.1	967	12	BM164469	BM164469 AGENCOURT
29	1105	59.8	952	9	AL539237	AL539237 AGENCOURT
30	1100	59.1	973	9	AL547220	AL547220 AGENCOURT
31	1099	59.7	776	10	AV725392	AV725392 AGENCOURT
32	1096.5	59.6	685	14	BM742526	BM742526 AGENCOURT
33	1096.5	59.6	685	14	BM672095	BM672095 AGENCOURT
34	1095	59.5	955	14	BQ172640	BQ172640 AGENCOURT
35	1092.5	59.4	638	13	BI088605	BI088605 AGENCOURT
36	1088	59.1	664	13	BI596603	BI596603 AGENCOURT
37	1084	58.9	956	9	AL566260	AL566260 AGENCOURT
38	1079	58.6	655	13	BI112632	BI112632 AGENCOURT
39	1074	58.4	675	10	BE258152	BE258152 AGENCOURT
40	1071.5	58.2	861	13	BI550693	BI550693 AGENCOURT
41	1066	57.9	633	13	BI792000	BI792000 AGENCOURT
42	1065	57.9	904	12	BE684268	BE684268 AGENCOURT
43	1058.5	57.5	705	10	BE544830	BE544830 AGENCOURT
44	1057.5	57.5	684	10	AV723717	AV723717 AGENCOURT
45	1056.5	57.4	1041	13	BM563365	BM563365 AGENCOURT

ALIGNMENTS

RESULT 1
AF161496
LOCUS AF161496 1860 bp mRNA linear HTC 22-MAY-2001
DEFINITION Homo sapiens HSPC147 mRNA, complete cds.
ACCESSION AF161496
VERSION AF161496.1 GI:6841515
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1860)
Zhang,Q.H., Ye,M., Wu,X.Y., Ren,S.X., Zhao,M., Zhao,C.J., Fu,G.,
Shen,Y., Fan,H.Y., Lu,G., Zhong,M., Xu,X.R., Han,Z.G., Zhang,J.W.,
Tao,J., Huang,Q.H., Zhou,J., Hu,G.X., Gu,J., Chen,S.J. and Chen,Z.

REFERENCE

1 (bases 1 to 1860)
Zhang,Q.H., Ye,M., Wu,X.Y., Ren,S.X., Zhao,M., Zhao,C.J., Fu,G.,
Shen,Y., Fan,H.Y., Lu,G., Zhong,M., Xu,X.R., Han,Z.G., Zhang,J.W.,
Tao,J., Huang,Q.H., Zhou,J., Hu,G.X., Gu,J., Chen,S.J. and Chen,Z.

TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)
MEDLINE	20499367
PUBMED	11042152
REFERENCE	
AUTHORS	1 (bases 1 to 1860) 2 (bases 1 to 1860) Ye, M., Zhang, Q. H., Zhou, J., Shen, Y., Wu, X. Y., Guan, Z. Q., Wang, L., Fan, H. Y., Mao, Y. F., Dai, M., Huang, Q. H., Chen, S. J. and Chen, Z.
TITLE	Human full length cDNA cloned from cd34+ stem cells
JOURNAL	Unpublished
REFERENCE	
AUTHORS	3 (bases 1 to 1860) Ye, M., Zhang, Q. H., Zhou, J., Shen, Y., Wu, X. Y., Guan, Z. Q., Wang, L., Fan, H. Y., Mao, Y. F., Dai, M., Huang, Q. H., Chen, S. J. and Chen, Z.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUN-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China
FEATURES	Location/Qualifiers
source	1. 1860 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CBPBN08" /cell_type="cd34+ stem cells" /tissue_type="blood" 302. 1354 /codon_start=1 /product="HSPC147" /protein_id="AAP2911.1" /db_xref="GI:6841516" /translation="MAHQPTPLTSSGTRRPVVDLAFSGIPYGYELISACDKGEMLR QGDVDMGTFLGHRGAWGATLNKDKATPAADPTAKWDAVSGDELMTLAKHII VKYDVFQDSNYVLITGGQDKLRLIYDNLKPAPKEKISGHTSGIKALWSEDKQILIS ADYTVRLMDHATMEVKSILNFNMNSVEMLEYIPGELIVLYVGSIAFHSVSLDPLIK SEEPAPATINSALHPEKEFLVAGCEDSKLYVDYNSGHELESYGHGPIHCVRSPSD GEIVASGDEGTLRLMOTVGVKTYGMKCVLPREDSGELAPKIGPETTEEELEETIA SENSDCIPSPADVKA"
CDS	
BASE COUNT	521 a 390 c 425 g 524 t
ORIGIN	
Alignment Scores:	
Pred. No.:	8.5e-214 Length: 1860
Score:	1784.50 Matches: 338
Percent Similarity:	97.44% Conservative: 4
Best Local Similarity:	96.30% Mismatches: 8
Query Match:	96.96% Indels: 1
DB:	11 Gaps: 1
US-09-856-836-2 (1-351) x AF161496 (1-1860)	
QY	1 MetAlaMetArGlnGlnProLeuPProLysSerGlyHisThrArgProValAlaAsp*** 20
DB	302 ATGGCATGAGACAGCGCGCTCCTCTCTGCGCCACACCGACCGCTGTTGATTTG 361
QY	21 AlaPheSerGlyTLeuPProTyrGlyTyrPheLeuLleSerAlaCysLysAspGlyLys 40
DB	362 GCCTTCAGTGGCATTCACGCGCTTATGGGTATTTCTTAATCAGCGCTTCAAGAGATGCTAA 421
QY	41 PrometLeuArgGlnGlyAspThrGlyAspTyrPheGlyThrPheLeuGlyHisLysGly 60
DB	422 CCTATCTCTACGCCAGGAGATACAGGAGCTGGATTGGAACTTTTGGGTCTTAAAGG 481
QY	61 AlaValTrrGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
DB	482 GCTGTTGGGGGTCAACACGATTAAGATGCCACCAAGACAGCTACAGAGCTGCAGAT 541
QY	81 PheThrAlaLysValTrrAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB	542 TTCACAGCCAAATGTGGGATGCTGCTCTCAGGAGATGAATGATGACCCCGCTCATAA 601
QY	101 HisIleValLysThrValAspPheThrGlnAspSerAsnTyrLeu***TrrGlyGlyGln 120
DB	602 CAAATTGTCAAGACTGTGATTTCACGACGAGATAGTATTAATTATTGTTAACGGGGAGAC 661

QY	121	Asp1y5tLeuLeuAag11eTyrAspLeuAa5nlyProGlu1aGluProLy5tGlu11leSer	140
Db	662	GATAAACGTTTACCATAATATAGCTTGAAACAACCTTAAGCAGAACTTAAGGAATTAGT	721
QY	141	GlyH5tHrSerGly11eLysLysAlaLeuTrrPy5SerAspAp1y5g11leLeuSer	160
Db	722	GGTCATTCTTCTGGGATTAATAAAAAAGCTCTGTGGTGCAGTGAGATAACAGATCTTCT	781
QY	161	AlaaspAp1y5tHrValArgLeuTrrAspHisAlaHrMetHrGluValLysSerLeu	180
Db	782	GCTGTATGCAAAACGTTGCGACTTGGATATATCTACTAGACAGAAAGAAATCTCTA	841
QY	181	AspH5anMetSerValSerSerMetLutTrrLeProGluGlu11leLeuVal11le	200
Db	842	AATTTTAATATGTCGTGTAGTAGATGAATATATCTCTGAGGGAGAGATTGGTTATA	901
QY	201	ThrTrrGlyArgSer11leAlaPheHisSerAlaValSerLeuGluPro11leLysSerPhe	220
Db	902	ACTTATGGACGATCTATGCTTTTCATATGTCAGATAAGTTGGACCCAAATTAATCCCTT	961
QY	221	GluAlaProAlaThr11leAsnSerLaseSer**HisProGluLysGluPheLeuValAla	240
Db	962	GAAAGTCCTGCACACATCAATTCGCACTCTTCACTGCGAAGAAATTTCTTGTTGCA	1021
QY	241	GlyGlyLysLysPhePheLysLeuTrrLysTrrAspTrrAsnSerGlyGluGluLeuGluSer	260
Db	1022	GCGCGTGAAAGATTTTAAACTTTATAGATATGTTATATAGCGGAAGAATTTGAATCC	1081
QY	261	TyrLysGlyHisPheGlyPro11leHisCysValArgPheSerProAspGlyGluLeuTyr	280
Db	1082	TTCAAAGGACACTTGTGCTCTATTCACGTGTGATTAATTAAGTATTAAGTCGAAGGAACTCTAT	1141
QY	281	AlaSerGlySerGlyAspGlyHrLeuArgLeuTrrGlnTrrValValGlyLysThrTyr	300
Db	1142	GCCAGTGGTTCCAGAAAGATGGAACATTGAGACATRGCAAACTGTGTTGGAATAACGAT	1201
QY	301	GlyLeuTrrLysCysVal**ProGluGluAspSerGlyGluLeuAlaLysProLys11le	320
Db	1202	GCGCTTTGAAATGCTGCTCTCTCAAGAAAGTACTGTGAGCTCGCAAGCCAAAGATT	1261
QY	321	GlyPheProGluThrAlaGluGluGluLeuAlaGluGlu11leAlaSerGlyLysnSerAsp	340
Db	1262	GGTTTTCACAGACACACAAAGAGAGGAGCTA--GAAAGAAATTGCTTCAGAGATTCAGAT	1318
QY	341	Ser11eTrrSerSerThrProGluValLysAla 351	
Db	1319	TGCATCTTCTCCTTCACGCTGATGTTAAGGCC 1351	
RESULT 2			
AUI23736		904 bp	mRNA
LOCUS	AUI23736		linear
DEFINITION	AUI23736 NT2RM2 Homo sapiens cDNA clone NT2RM200935 5', mRNA		EST 01-AUG-2002
ACCESSION	AUI23736		
VERSION	AUI23736.1		GI:10948452
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 904)		
AUTHORS	Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Isogai,T.		
TITLE	HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S., Isogai,T.)		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan		

Tel: 81-438-52-3975
Fax: 81-438-52-3986

Email: genomics@hri.co.jp
HRI human cDNA project: 5'-6' end one pass sequencing: Helix
Research Institute: cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

source

Location/Qualifiers

1..904

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="NT2RM200935"

/clone_lib="NT2RM2"

/cell_type="teratocarcinoma"

/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"

BASE COUNT 257 a 173 c 214 g 256 t 4 others

ORIGIN

Alignment Scores:

Pred. No.: 1,54e-168 Length: 904
Score: 1422.00 Matches: 279
Percent Similarity: 95.59% Conservative: 3
Best Local Similarity: 94.58% Mismatches: 11
Query Match: 77.28% Indels: 3
DB: 9 Gaps: 0

US-09-856-836-2 (1-351) x AU123736 (1-904)

```

OY 1 MetAlaMetArgInuThrProLeuThCysSerGlyHisThrArgProValAlaASP*** 20
    |||||||
DB 24 ATGCATATGACACAGACGCGCTACCTGCTGGCCACGCGACCGCGTGGTATTG 83
OY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
    |||||||
DB 84 GCCCTTCAGTGGCAGTACGCGCTTATGGTATTTCTTATTCAGCGCTTCAAGATGTTAA 143
OY 41 ProMetLeuArgGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
    |||||||
DB 144 CTTATGCTACGCGCAGGAGATACGAGACTGATGGAATTTTGGGTCAATTAAGT 203
OY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaASP 80
    |||||||
DB 204 GCTGTTGGGGTGAACACGATTAAGATGCCACCAACACTTACGACGCTGGAGAT 263
OY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
    |||||||
DB 264 TTCACAGCCAAAGTGGGATGCTGCTCAGAGATGAATGATGACCCGCTCATTA 323
OY 101 HisIleValLysThrValAspPheThrGlnAspSerAspTyrLeu***ThrGlyGlyGln 120
    |||||||
DB 324 CACATATGTCAGACTGTGATTTTCACGACGATAGTAATATTGTTAACCGGGGACAG 383
OY 121 AspLysLeuLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluIleSer 140
    |||||||
DB 384 GATTAACCTGTTCGATATATGACTTGAACAAACCTGAAGCAGAACTTAAGAAATTA 443
OY 141 GlyHisThrSerGlyIleLysLysAlaLeuTrpCysSerAspAspLysGluIleLeuSer 160
    |||||||
DB 444 GGTCACTTCTGGTATTAATAAAAAAAGCTGTGTCAGAGAGATTAACAGATTTCTTCT 503
OY 161 AlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeu 180
    |||||||
DB 504 GCGATGACAAAACCTGTGCTGGATTCATGCTACATATGACAGAACTGAATCTCTA 563
OY 181 AsnPhaAsnMetSerValSerSerMetGluTyrIleProGluGlyGluIleLeuValIle 200
    |||||||
DB 564 AATTTTAATATGCTGTTGATGATGAGATATATTCCTGAGAGAGAGATTTGGTTATA 623
OY 201 ThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProLysSerPhe 220
    |||||||
DB 624 ACTTATGAGCAGATCTATTCCTTTTCACTAGTGCAGTAAGTTGGACCAATTAATCCCTTT 683

```

```

OY 221 GluAlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAla 240
    |||||||
DB 684 GAACTCTCTGACACCAATTCATCTCTCATCTCTGCTGGAATTAATTTCTTGGTGA 743
OY 241 GlyGlyGluAspPheLysLeuTyrLysTyrAspPyr-AsnSerGlyGluGluLeuGlu 260
    |||||||
DB 744 NGCGGTGAACATTTTAACCTTTAATAGTATGATTAATTAATAGTGAACAAATTAAGATC 803
OY 260 rTyrLysGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyGluLeu 280
    |||||||
DB 804 CTC-AAGGACACTTGTGCTCTATTCACCTGGTGAGATTTTAATCCATGAGAACTCTA 862
OY 280 rAlaSerGlySerGluAsp-GlyThrLeuArgLeuTrpGln 293
    |||||||
DB 863 TGCCAGNGGTTCAAAAGATTTGGAACATTTGGACTNTGGCAA 903

RESULT 3
BM802585
LOCUS
DEFINITION
  BM802585 1174 bp mRNA linear EST 05-MAR-2002
  5' mRNA sequence.
ACCESSION
  BM802585
VERSION
  BM802585.1 GI:19119408
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
  1 (bases 1 to 1174)
  NIH-MGC http://mgi.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-remail.nih.gov
  Tissue Procurement: ATCC
  CDNA Library Preparation: Life Technologies, Inc.
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNI)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LMNI at:
  http://image.llnl.gov
  Plate: LAM12324 row: b column: 24
  High quality sequence stop: 656.

FEATURES
  source
    1..1174
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:5574839"
    /clone_lib="NIH-MGC_92"
    /tissue_type="embryonal carcinoma, cell line"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: testis; Vector: PCMV-SPORT6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
    Average insert size 2.5 kb. Library enriched for
    full-length clones and constructed by Life Technologies.
    Note: this is a NIH-MGC library."

BASE COUNT 369 a 232 c 257 g 316 t

ORIGIN

Alignment Scores:
Pred. No.: 4.29e-163 Length: 1174
Score: 1380.50 Matches: 288
Percent Similarity: 88.46% Conservative: 11
Best Local Similarity: 85.21% Mismatches: 27
Query Match: 75.03% Indels: 12
DB: 14 Gaps: 6

US-09-856-836-2 (1-351) x BM802585 (1-1174)

OY 22 PheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLysPro 41
    |||||||
DB 3 TTCAGTGGCAGTACGCGCTTATGGTATTTCTTATTCAGCGCTTGAACAAAGATGAACCT 62

```

QY 42 MetIeuArgInGlnIAspThrGlnIAspThrIleGlyThrPheLeuGlyHisLysGlyAla 61
 Db 63 ATCTGACGCGAGGAGATACAGAGACTGATGGACATTTTGGGTCAATAAGAGTCT 122
 QY 62 ValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaIleThrAlaIleAspPhe 81
 Db 123 GTTGGGCTGCAACACAGATTAAGATGCCAACAGCAGCTACAGCAGCATTC 182
 QY 82 ThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLysHis 101
 Db 183 ACACGCCAAATGTGGGATGCTGCTCTCAGAGATGAATGATGACCTGGCTCATAAACAC 242
 QY 102 IleValLysThrValAspPheThrGlnAspSerAsnTrpLeu***ThrGlyGlyAsp 121
 Db 243 ATGTCTAAGACTGTGATTCACCGCAGAGATGTAATATTGTTAAACCGGGGAGAGAT 302
 QY 122 LysLeuLeuArgIleLeuTrpAspLeuAsnLysProGluIleGluProLysGluIleSerGly 141
 Db 303 AAAGCTTTAGCATATATGACTTCAACAACTGACAGACAGAACTTAAGAAATAGTCT 362
 QY 142 HisThrSerGlyIleLysLysAlaLeuTrpCysSerAspLysGluIleLeuSerAla 161
 Db 363 CATACTCTGGATATAAAAAAGCTGTGTCAGTCAGATGAATAAACAGATTCTTCTGCT 422
 QY 162 AspAspLysThrValAspGluLeuTrpAspHisAlaThrMetThrGluValLysSerLeuAsn 181
 Db 423 GATACCAAACTGTTCGACTTGGATCATCTACTATGACAGAAAGTAAATCTCTAAT 482
 QY 182 PheAsnMetSerValSerSerMetGluTrpIleProGluGlyGluIleLeuValIleThr 201
 Db 483 TTTATATAGCTGTAGTAGATGATGAATATATCTCGAGGAGAAATTTGGTTATTAAT 542
 QY 202 TyrGlyAspSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPheGlu 221
 Db 543 TATGACACATCTATGCTTTTCATAGTCAGTAACTTGGACCCCAATTAATCTTTGAA 602
 QY 222 AlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAlaGly 241
 Db 603 GCTCTGACACATCAATCTGCATCTCTCATCTGAGAAAGATTTCTTCTGACAGC 662
 QY 242 GlyLysAspPheLysLeuTrpLysTrpAspTrpAsnSerGlyGluLeuGluSerTrp 261
 Db 663 GGTGAAGATTATTAACCTTATAGTATGATTAATAGTGAAGAAATTAACATCTTAC 722
 QY 262 LysGlyHisPheGlyProIleHisCys-ValArgPheSer-ProAspGlyGluLeu-Tyr 280
 Db 723 AAGGACACTTGTGCTATTCACCTGGGTGAGATTTAGTCCCGATGAGAACATCTTAT 782
 QY 281 AlaSerGly---SerGluAspGly-ThrLeuArgLeuTrpGlnThrValVal---GlyLys 298
 Db 783 GCCATGGGTTTCAGAAAGATGGAACATTTGAGACTATGCAAAACTGTGGGTAGAGAAA 842
 QY 298 sThrTrpGlyLeuTrpLysCysVal***-ProGluGluAsp-----SerGlyLys 315
 Db 843 AACGATGGCCCTTTGAAATATGTCCTTCCCGGAAAAAAGTGGGTGGAGACCTGC 902
 QY 315 uAlaLysProLysIleGlyPhe---ProGluThrAlaGlu-GluGluLeuAlaGluLys 334
 Db 903 CACAACCCCAAAAGATTTGCTTCCCGGAAACACCAAGAGGAGCACTTTAAAGAAA 962
 QY 334 lAlaLysGluAsn---SerAspSerIleTrpSerSerThrPro 347
 Db 963 TTGCTTCCAAAAAATTTGCAATGAGCATCTTTCTTCCAGCCT 1006
 RESULT 4
 AL517952 944 bp mRNA linear EST 13-FEB-2001
 LOCUS AL517952 LTI_NFLO11.NBC1 Homo sapiens cDNA clone CS0DA004YH09 5
 DEFINITION AL517952 LTI_NFLO11.NBC1 Homo sapiens cDNA clone CS0DA004YH09 5
 prime, mRNA sequence.
 ACCESSION AL517952
 VERSION AL517952.1 GI:12781445
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 944)
 AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 FEATURES
 source
 1..944
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DA004YH09"
 /clone_id="LTI_NFLO11.NBC1"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
 BASE COUNT 259 a 197 c 235 g 250 t 3 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 02e-156 Length: 944
 Score: 1325.00 Matches: 256
 Percent Similarity: 98.10% Conservative: 2
 Best Local Similarity: 97.34% Mismatches: 5
 Query Match: 92.01% Indels: 2
 DB: 9 Gaps: 0
 US-09-856-836-2 (1-351) x AL517952 (1-944)
 QY 1 MetAlaMetArgGlnThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
 Db 158 ATGGCAATGAGACAGACGCCGCTCCTGCTGGCCACACGACCGCTGGTGAATTTG 217
 QY 21 AlaPheSerGlyTrpLeuProGlyTrpPheLeuIleSerAlaCysLysAspGlyLys 40
 Db 218 GCCCTGAGTGGCATCGCGCTTATGGGATTTCTTATATGACGCCCTTGCAAAAGTGTAA 277
 QY 41 PrometLeuArgGlnIAspThrGlnIAspTrpIleGlyThrPheLeuGlyHisLysGly 60
 Db 278 CCATCTACGCCAGGAGATACAGAGACTGATTTGAAACATTTTGGGTCAATAAGCT 337
 QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaIleThrAlaIleAsp 80
 Db 338 GCTGTGGGCTGCAACACTGATTAAGATGCCACCAAGCAGCTACAGCAGCTGAGAT 397
 QY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
 Db 398 TTACAGCCCAAGATGTGGATGCTGTCTCAGAGATGAATGATGACCGCTGCTATAA 457
 QY 101 HisIleValLysThrValAspPheThrGlnAspSerAsnTrpLeu***ThrGlyGlyGln 120
 Db 458 CACATTTGTCAAGACTGTGATTTCCACGACAGATGTAATATTGTTAAACCGGGGACAG 517
 QY 121 AspLysLeuLeuArgIleTrpAspLeuAsnLysProGluIleGluProLysGluIleSer 140
 Db 518 GATTAACCTTTAGCATATATGACTTGAACAAACCTGAAACAGAACTTAAGAAATTAGT 577

QY 141 GlyHisThrSerGlyIleLysLysAlaLeuTrpCysSerAspAspLysGlnIleLeuSer 160
 Db 578 GGTATCTCTGTTGTAATAAAAAAGCTGTGTGAGTGAAGATAAAGATCTCTTCT 637
 QY 161 AlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetTrpGluValLysSerLeu 180
 Db 638 GCTGATACAAAAGTCTTCCACTTGGGATCATGCTACTATGACAGAAAGTAATCTCTA 697
 QY 181 AspPheAsnMetSerValSerSerMetGluTrpIleProGluGluIleLeuValIle 200
 Db 698 AATTTTAAATGCTGTTAGTATGATGATATATCTCTAGGAGAGATTTGGTTATA 757
 QY 201 ThrTyrgLysArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe 220
 Db 758 ACTATGAGACGATCTATGCTTTTCATATGACATGAGTTGGACCAATTAATCTCTT 817
 QY 221 GluAlaProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuValAla 240
 Db 818 GAAGCTCTGCAACACATCTTCATCTTCATCTTCATCTTCATCTTCATCTTCATCTTCAT 877
 QY 241 GlyGlyGluAspPheLysLeuTrpLysTyrgAspTrpAsnSerGlyLysGluLeuGluSer 260
 Db 878 GGGGGGTAAGATTTTAACTTTATAGTATGATTAATAGTGGGAGRA-TTAGAATC- 935
 QY 261 TyrlsGly 263
 Db 936 TACAAGGG 944

RESULT 5
 BM472016 1055 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT 6464560 NIH_MGC_67 Homo sapiens cdna clone IMAGE:5581667
 DEFINITION 5', mRNA sequence.
 ACCESSION BM472016
 VERSION BM472016.1 GI:18521058
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1055)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs.fda.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: LMNL2341 row: 0 column: 12
 High quality sequence start: 3
 High quality sequence stop: 658.
 Location/Qualifiers
 1. 1055
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5581667"
 /clone_id="NIH_MGC_67"
 /lssue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

FEATURES
 source
 BASE COUNT 308 a 222 c 247 g 277 t 1 others
 ORIGIN
 Alignment Scores: 1.04e-154 Length: 1055
 Pred. No.:

Score: 1313.50 Matches: 268
 Percent Similarity: 87.74% Conservative: 11
 Best Local Similarity: 84.28% Mismatches: 29
 Query Match: 71.39% Indels: 10
 DB: 13 Gaps: 4
 US-09-856-836-2 (1-351) x BM472016 (1-1055)

QY 1 MetalMetArgGlnThrProLeuThrCysSerGlyHisThrArgProValAlaAsp** 20
 Db 68 ATGGCAATGAGACAGACGCCCTCACTCTCTGGCACACGCCGCGTGTATTTG 127
 QY 21 AlaPheSerGlyIleThrProTrpGlyTrpPheLeuIleSerAlaCysLysAspLys 40
 Db 128 GCCTTACGAGCATCAGCGCTTATGGGTATTTCTTAATCAGCGCTTGCMAAGATGTAA 187
 QY 41 PrometLeuArgGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
 Db 188 CCTATGCTACGCGCAGAGATACAGAGAGATTTGGACATTTTGGTCTATTAAGT 247
 QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaTrpLysAlaAlaThrAlaAlaAsp 80
 Db 248 GCTGTTGGGGTGCACACCTGAAATAGATGACCAAAAGAGCTACAGACCTGCAGAT 307
 QY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
 Db 308 TTCACAGCCAAAGTGTGGATCTCTCTCAGAGATGATGATGATGATGATGATGATGAT 367
 QY 101 HisIleValLysThrValAspPheThrGlnAspSerAsnTrpLeu**ThrGlyGlyGln 120
 Db 368 CACATTTGCAACAGCTGTGATTTCCAGCAGATGATTAATTTGTTAAACCGGGGACAG 427
 QY 121 AspLysLeuLeuArgIleThrAspLeuAsnLysProGluAlaGluProLysGluIleSer 140
 Db 428 GATTAACCTTTACGATATATGACTTGAACCAACCTGAAAGCAACCTTAAGAAATTAAGT 487
 QY 141 GlyHisThrSerGlyIleLysLysAlaLeuTrpCysSerAspAspLysGlnIleLeuSer 160
 Db 488 GGTCTACTCTGTGTAATAAAAAAGCTGTGTGAGTGAAGATAAAGATCTTCT 547
 QY 161 AlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetTrpGluValLysSerLeu 180
 Db 548 GCTGATGACAAAAGCTGTGACCTTTGGATCATCTCTATGACAGAAAGTAATCTCTA 607
 QY 181 AspPheAsnMetSerValSerSerMetGluTrpIleProGluGluIleLeuValIle 200
 Db 608 AATTTTAAATGCTGTTAGTATGATGATATATCTCTGAGGAGAGATTTGGTTATA 667
 QY 201 ThrTyrgLysArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe 220
 Db 668 ACTATGAGACATCTATGCTTTTCATATGAGTACAGTAGTGGACCCCATTAATCTCTT 727
 QY 221 GluAlaProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuValAla 240
 Db 728 GAAGCTCTGCAACACATCTTCATCTTCATCTTCATCTTCATCTTCATCTTCATCTTCAT 787
 QY 241 GlyLysLysAsp-PheLysLeuTrpLysTyrgAspTrpAsnSerGlyLysGluLeuGluSer 260
 Db 788 GCCGCGCAAGAAATTTTACCTTATATGATGATTAATATAGGAGGAGAAAGATTTTACA 847
 QY 259 uSerTyrgLysGlyHis-PheGlyProIleHisCys-ValArgPheSerProAspLysGly 278
 Db 848 TTCTTACAGGAGGACCTTGGGCTTAACTGCTGAGTAAATTTAGTTCCTGCGGCGGA 907
 QY 278 u---LeuTyrgAla-SerGlySerGluAspLysThrLeuArgLeuTrpGlnThrValVal- 296
 Db 908 AAAGCTCTATGCCATGAGGTTCCTCCAGAAAGCAACATTTGCAACTATGCGCAACATCTGG 967
 QY 297 -----GlyLysThrTyrgLysLeuTrpLys---CysVal**Pro 308
 Db 968 AACGCAAAAAGCTATGAGCTTTGGAATAATAGTGTGCTCTCT 1011

RESULT 6

Directionally cloned using the following adaptors:

5'-TCGACCCAGGCTCCG-3' and
5'-GACTAGTTCAGATCGGAGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lipski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies.

BASE COUNT 257 a 186 c 223 g 236 t
ORIGIN

Alignment Scores:

Pred. No.: 5,95e-149 Length: 902
Score: 1267.00 Matches: 248
Percent Similarity: 95.47% Conservative: 5
Best Local Similarity: 93.58% Mismatches: 7
Query Match: 68.86% Indels: 5
DB: 14 Gaps: 1

US-09-856-836-2 (1-351) x BQ889506 (1-902)

QY 1 MetAlaMetAArgInThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
DB 115 ATGGCAATGAGACAGACGCGCTCAGCTGCTGGCCACAGCGACCGCTGATTG 174
QY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 175 GCCCTTACGTGACATCAGCGCTTATGGTATTTCTTAATCAGCGCTTCAAGAGTGTAA 234
QY 41 PrometLeuArgInGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
DB 235 CCTATGCTACGCGACGAGATACAGAGACTGATGTAACATTTTGGGTCAATAAGT 294
QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
DB 295 GCTGTGGGCTGCAACACTGAATAGATGCCACCAACACCTACAGACGCTGCAGAT 354
QY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 355 TTCACAGCCAAAGTGGATGGCTGCTCAGAGATGATGATGACCTCGCTCATAAA 414
QY 101 HisIleAlaLysThrValAspPheThrGlnAspSerAsnTyrLeu***ThrGlyGlyGln 120
DB 415 CCAATGCTCAAGACTGTGATTCACGCGAGATGAAATATTGTTAAACGGGGACAG 474
QY 121 AspLysLeuLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluIleSer 140
DB 475 GATTAACGTGTACGATATATGACTTGAACAAACCTGAAGCAGAACCTTAAGAAATTAGT 534
QY 141 GlyHisThrSerGlyIleLysLysAlaLeuTrpCysSerAspAspLysGlnIleLeuSer 160
DB 535 GGTCAACTCTCTGTATTAATAAAAGCTGTGTGTGAGAGATGAACAGATTCTTCT 594
QY 161 AlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeu 180
DB 595 GCGATGACAAAACCTGTGACTTGGATCATGCTACTATGACAAAGTGAATCTCTCA 654
QY 181 AsnPheAsnMetSerValSerSerMetGluTyrIleProGluGlyLysLeuValIle 200
DB 655 AATTAAATATGCTGTATGATGATGATATATTCCTGAGAGAGACAGATTGGTATTA 714
QY 201 ThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe 220
DB 715 ACTTATGAGAGATCTAATCTTTCTTATGTCAGTAAGTTTGACCCATTTAATCCTTT 774
QY 221 GluAlaProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAla 240
DB 775 GAAGCTCCGCAACCATCTCTGATCTTCATCCGAGAAAGAAATTTCTTGTGCA 834
QY 241 Gly--GlyGluAspPheLysLeuTyrLysTyr--AspTyrAsnSerGlyGluLysLeu 259
DB 835 GCGCGGGGAAATTTTAACTTTATAGTATGATTAATATAGTGCAGAAAGATTT--- 891

QY 260 SerTyrLys 262
DB 892 ---TACAAA 897

RESULT 8
BQ439153
LOCUS

DEFINITION BQ439153 875 bp mRNA linear EST 24-MAY-2002

ACCESSION ACENCOUPT_7827244 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014262

VERSION BQ439153.1 GI:21178229

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 875)
NIH-MGC <http://mgi.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Place: LLNL3208 row: h column: 07
High quality sequence stop: 675.

FEATURES
location/Qualifiers
1..875
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6014262"
/clone="11b-NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-Sport6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: This is a NIH-MGC Library."

BASE COUNT 253 a 175 c 209 g 235 t 3 others
ORIGIN

Alignment Scores:

Pred. No.: 1.01e-148 Length: 875
Score: 1265.00 Matches: 248
Percent Similarity: 95.80% Conservative: 3
Best Local Similarity: 94.66% Mismatches: 8
Query Match: 68.75% Indels: 3
DB: 14 Gaps: 0

US-09-856-836-2 (1-351) x BQ439153 (1-875)

QY 1 MetAlaMetAArgInThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
DB 90 ATGGCAATGAGACAGACGCGCTCAGCTGCTGGCCACAGCGACCGTGTGATTG 149
QY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 150 GCTTCACTGACATCAGCGCTTATGGGATTTCTTAATCAGCGCTTCAAGAGTGTAA 209
QY 41 PrometLeuArgInGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
DB 210 CCTATGCTACGCGACGAGATACAGAGACTGATGTAACATTTTGGGTCAATAAGT 269
QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
DB 270 GCTGTGGGCTGCAACACTGAATTAAGATGCCACCAAGACGCTTACAGACGCTGCAGAT 329

OY	81	PheThrAlaIysValITrPaspAlaValSerCylspsGluMetThrLeuAlaHisLys	100
Db	330	TTTCACAGCCAAAGTGTGGATGTCTGTCCAGAGATGATTAATGATGACCTGGCTCATATAA	389
OY	101	HisIleValIysThrValAspPheThrGlnAspSerAsnIyrIleu**ThrGlyGln	120
Db	390	CACATTGTCAAGACGTGTGCATTTCACGACAGTAGTATTAATTTCTTTAAACCGGGGACAG	449
OY	121	AspLysLeuLeuAlaGlyIleTyrAspLeuAsnLysProGluAlaGluProLysGluIleSer	140
Db	450	GATAAACGTTCAGCATATATGACTTGAACAAACCTGGACAGAACCTTAAGCAATTTGCT	509
OY	141	GlyHisThrSerGlyIleLysLysAlaLeuThrPcySerAspAspLysGlnIleLeuSer	160
Db	510	GGTCACTCTTCTGGTATATAAAAGACCTGTGGTGCATGAGATTAACAGATTTCTTCT	569
OY	161	AlaAspAspLysThrValArgLeuThrPAspHisAlaThrMetThrGluValLysSerLeu	180
Db	570	GCTGATGACAAACAGCTGTGACCTTGGGGATCAAGCTACTATGACAGAAAGTGAATCTCTA	629
OY	181	AsnPheAsnMetSerValSerSerMetGluIyrIleProGluGluIleLeuValIle-	200
Db	630	AATTTAATATATCTCTGTATGTAGTATGTGAATATATCTCGAAGGAGAGATTTTGTGTATA	689
OY	201	ThrIyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSer-PH	220
Db	690	ACTTATGGACGAGTATATGCTTTCATACAGTACAGTAAGTTGGACCAATTATATCCNTT	749
OY	220	ecGluAlaProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuValAla	240
Db	750	TGAAGCTCTCGAACCATCAATTCATCTGCATCTCTCATCTCGAAGAAAGATTTCTGTCTGC	809
OY	240	agIlyGlyGluAspPhe-LysIleuTyrLysTyrAspLysAsnSerGly-GluGluLeuGlu	259
Db	810	AGCGGAGNAGATTTTAAACCTTATAGATATGATATATATAGTGGAGACAGATATAGAA	869

RESULT 9
 BG687473
 LOCUS
 DEFINITION
 602839445F1 NTH_MGC_59 Homo sapiens CDNA clone IMAGE:4762475 5',
 mRNA sequence.
 ACCESSION
 BG687473
 VERSION
 BG687473.1 GI:13918870
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 757)
 NIH-MGC <http://imgc.nci.nih.gov/>.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONETECH Laboratories, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1616 row: 3 column: 12
 High quality sequence stop: 723.

Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATATGACG-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCCGAGCGCCGACATG-drr(30)-BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH/MGC Library."

BASE COUNT	235 a	120 c	182 g	220 t
ORIGIN				
Alignment Scores:				
Pred. No.:	4.83e-145			Length: 75
Score:	1235.00			Matches: 23
Percent Similarity:	94.86%			Conservative: 2
Best Local Similarity:	94.07%			Mismatches: 1
Query Match:	67.12%			Indels: 12
DB:	12			Gaps: 0

US-09-856-836-2 (1-351) x BG687473 (1-757)

QY	77	ALAALAAIAASpRheThrAlaIysValTrpAspAlaValSerGlyAspGluLeuMetThr	96
Db	1	GCAGCTGCAGATTCACAGCCAAAGTGTGGAGTGTCTCAGGAAATATATTGATGACC	60
OY	97	LeuAlaHisLysHisLileValLysThrValAspPheThrGlnAspSerAsnTyrLeu**	116
Db	61	CTGGCTCTATAACACATTGTCTCAAGACTGTGGATTTCACGAGATAGTAATTAATTTGTGA	120
OY	117	ThrGlyGlyGlnAspLysLeuLeuArgLleTyrAspLeuAsnLysProGlnIleGluPro	136
Db	121	ACCGGGGACAGAGATAAACGTTCGAGATATACACTGAAACAAACCTGAAGACAACT	180
OY	137	LysGluIleSerGlyHisIsthSerGlyLleLysLysAlaLeuThrPcysSerAspLys	156
Db	181	AAGGAATTAATAGTCGCACTTCGTGGATATAAAAAAACCTCTGGTGGAGTACAGATAA	240
OY	157	GlnIleLeuSerAlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGlu	176
Db	241	CAGATTCCTTCGCGCATGACGACAAACCTGTTCGACTTTGGGATCATGCTACTATGACAGAA	300
OY	177	ValLysSerLeuAsnPheAsnMetSerValSerSerMetGluTyrLleProGluGlyGlu	196
Db	301	GGAAATCTCTAAATTTAAATATGTCTGTATGTATGGATATATATCTCGAGGAGAG	360
OY	197	IleLeuValIleThrTyrGlyArgSerLleAlaIleHisSerAlaValSerLeuGluPro	216
Db	361	ATTTCGTATTAACATTATGACGAGATCTATTCCTTTATCATGTCGACGTAACTTTGGACCA	420
OY	217	IleLysSerPheGluAlaProIleThrLleAsnSerAlaSer**HisProGluLysGlu	236
Db	421	ATTAAATCCCTTGAAGCTCTCGCAACCAATTCATTCGATCTCTTCATCTCGGACAGAA	480
OY	237	PheLeuValAlaGlyGlyGluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGlu	256
Db	481	TTTCTGTGTGACGGGGGGAAGATTTAACCTTATAAGTATGATTAATATAGCGAGAA	540
OY	257	GluLeuGluSerTyrLysGlyHisPheGlyProIleHisCysValAlaArgPheSerProAsp	276
Db	541	GAATTGGAATCCCTCAAGGACACTTTCGCTCATATCACTGTGGAATTAATTAAGTCCGAT	600
OY	277	GlyGluLeuTyrAlaSerGlySerGluAspGlyThrLeuArgLeuTyrPglInThrValVal	296
Db	601	GGGAAACTATGCCCACTGCTTCACAAAGATGCACATTTGAACATATGCGCAAACTGTGTA	660
OY	297	GlyLysThrTyrGlyLeuThrPlyScyVal**ProGluGluAspSerGlyGluLeu-Al	316
Db	661	GGAAAAACGATANGGCTTGACCATGTGCTCTTGAAGAAGATAGTGTGAGCTGGGC	720
OY	316	AlysProLysIleGlyPheProGluThrAlaGluGlu 328	

Db 721 AAGGCAAGATGCTATTCAGAGACAGACAGAG 757

RESULT 10
AJ395256 822 bp mRNA linear EST 25-JAN-2001

LOCUS AJ395256 dktz426 Gallus gallus cDNA clone 21m101, mRNA sequence.

DEFINITION AJ395256

ACCESSION AJ395256

VERSION AJ395256.1 GI:7125986

KEYWORDS EST.

SOURCE chicken.

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 822)

AUTHORS Adirakmanov, I., Lodygin, D., Gerolt, P., Arakawa, H., Law, A., Plachy, J., Korn, B., and Buerstedde, J.M.

TITLE A large database of chicken bursal ESTs as a resource for the analysis of vertebrate gene function

JOURNAL Genome Res. 10 (12), 2062-2069 (2000)

MEDLINE 20568495

COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.
Location/Qualifiers

FEATURES
source
1. 822
/organism="Gallus gallus"
/strain="CB"
/db_xref="taxon:9031"
/clone_1lb="21m101"
/clone_1lb="dktz426"
/issue_type="Bursa of Fabricius"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"

BASE COUNT 249 a 150 c 211 g 211 t 1 others

ORIGIN

Alignment Scores:
Pred. No.: 3 7e-144 Length: 822
Score: 1228.50 Matches: 238
Percent Similarity: 92.34% Conservative: 15
Best Local Similarity: 86.86% Mismatches: 18
Query Match: 66.77% Indels: 3
DB: 9 Gaps: 1

US-09-856-836-2 (1-351) x AJ395256 (1-822)

QY 52 IleglYThrPheLengUlyHsIysGlyAlaValTTPGlyAlaThrlenuAsnLysAspAla 71
|||||
Db 3 ATGGGACGTTCTAGTGTATTAAGTGTCTGCTGGGGTCTACTTTGACAAAGATGCC 62
|||||
QY 72 ThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTTPAspAlaValSergly 91
|||||
Db 63 ACTAAAGGACGACGAGGAGCTGACATTTTACAGCAAAAGTGGGATGCTGTGCGGA 122
|||||
QY 92 AspGluLeuMetThrLeuAlaHisLysHisLysValLysValAspPheThrLinsp 111
|||||
Db 123 GATGAACATACATACATTTGGCTTACAAACATTTGCAAAAGTGGATTTTACACAGAT 182
|||||
QY 112 SerAsnTyrLeu**ThrGlyGlyLysAspLysLeuLeuArgLysTyrAspLysLys 131
|||||
Db 183 AGCAATTAATCTGTAAACAGGTGACACAGATTAATGTTCGCGATCTATGATTTAAGCAG 242
|||||
QY 132 ProGluAlaGluProLysGluLysSerGlyHisThrSerglyLysLysAlaLeuTrp 151
|||||
Db 243 CCGGAAGCAGAACCTGATCTGTGACGTGGCATACTTCTGTATTAAAGAGCTTTATGC 302
|||||
QY 152 CysSerAspAspLysGlnLysLeuSerAlaAspAspLysThrValArgLeuTrpAspHis 171
|||||
Db 303 AGCAGTGATGATTAACAGATCTTTACACCTGATGATTAAGTGTCCCTCTGGAGCAG 362
|||||

QY 172 AlaThrMetThrGluValLysSergLeuAsnPhasMetSergValSergMetGluTyr 191
:::|||||
Db 363 AGTACATGACTGAAAGGAGGACCTAAATGTTGCTGTCTGTGACGACATGAGAT 422
|||||
QY 192 IleProGluGlyLysLeuValIleThrTyrGlyArgSergLysLeuAlaHisSera 211
:::|||||
Db 423 GTTCCAGAAAGGCAATTAATCTGTGATTAACCTTAAGCAACATTAATCTTTCATAGCGG 482
|||||
QY 212 ValSergLeuGluProLysSergPheGluAlaProAlaThrLysAsnSeraLaser 231
:::|||||
Db 483 GAAACCTTAAGCAAGATTAATTAATTTGAAGCAGCTGTACATCAATTCATCTGATCCCT 542
|||||
QY 232 HisProGluLysGluProLysLeuValAlaGlyGlyLysAspPheLysLeuTyrTyrAsp 251
|||||
Db 543 CACCCAGAGAAAGAAATGTTGTTGTCAGAGTGGGAGATTTAACTCTATAAATATGAC 602
|||||
QY 252 TyrAsnSergLysGlyLysLeuGluSergTyrLysGlyHisPheGlyProLysSera 271
|||||
Db 603 TATAACACAGAGAAAGCAATCAATCTTAACAAAGGCACTTTGTCATTAATCTGCTGTG 662
|||||
QY 272 ArgPheSergProAspGlyLysLeuTyrAlaSergLysSergLysAspGlyThr-LeuArgLe 291
|||||
Db 663 AGATTAGCCCTGACGAGGAGTATATGCAATGCTGCTGAGAGATGTAACCTTAAGCT 722
|||||
QY 291 UTTPGlnThrValAlaGlyLysThrTyrGlyLeuTrp-LysCysVal**ProGluGlu 311
|||||
Db 723 GGGGAGAGAAACAGTAAAGGAAACCTATGCTTTGTTGGAGAGTGTACTTCTGTAAGAG 782
|||||
QY 311 sPserGlyGluLeuAlaLysProLysIleGlyPhePro 323
:::|||||
Db 783 AGAATGACAGAA--AGCAGNCACAAAGCCAGGAGACCC 817
|||||

RESULT 11
BM450309 1019 bp mRNA linear EST 05-FEB-2002

LOCUS BM450309

DEFINITION AGENCOURT_6394168 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494620

ACCESSION BM450309

VERSION BM450309.1 GI:18499349

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1019)

NIH-MGC <http://mhc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

UNPUBLISHED (1999)

CONTACT: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LRAM12120 row: d column: 13

High quality sequence stop: 705.

Location/Qualifiers

1. 1019
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:5494620"
/clone_1lb="NIH_MGC_67"
/issue_type="retinoblastoma"
/lab_host="DH10b (phage-resistant)"
/note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Library constructed by Life Technologies."

BASE COUNT 281 a 215 c 257 g 263 t 3 others

ORIGIN

Alignment Scores:

Pred. No.: 1.33e-142 Length: 1019
 Score: 1217.50 Matches: 250
 Percent Similarity: 89.35% Conservative: 10
 Best Local Similarity: 85.91% Mismatches: 21
 Query Match: 66.17% Indels: 10
 DB: 13 Gaps: 4

US-09-856-836-2 (1-351) x BM450309 (1-1019)

```

QY 1 MetAlaMetArgGlnThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
DB 161 ATGGCAATGAGACAGACGCGCTACCTGCTGCGCACACGCCAGCCGCGTGGATTGG 220
QY 21 AlaPheSerGlyLeuThrProGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 221 GCCTTGAGTGCATCAGCGCTTATGGGTATTTCTTAATCAGCGCTGCAAGATGTAA 280
QY 41 PrometLeuArgGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
DB 281 CCTATGCTACGCGCAGGAGATACAGGAGACTGGATGGAACTTTGGGTCATTAAGGT 340
QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaAlaAlaAsp 80
DB 341 GCTGTTGGGGGTGACACACTGATTAAGGATGCCAACAAAGCAGTACAGCGTGCAGAT 400
QY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 401 TTCACACCCAAAGTGGGATGCTGCTCAGAGAGATGAATGATGAGACCTGCTCATTA 460
QY 101 HisIleValLysThrValAspPheThrGlnAspSerAsnTyrLeu**ThrGlyGlyGln 120
DB 461 CACATTTGCAAGACTGCTGATTTTCCAGCAGCATAGTATTAATTTTGTAAACCGGGGAG 520
QY 121 AspLysLeuLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluLysSer 140
DB 521 GATTAACCTGATACGATATATGACTTGAACAACTGAGACGAACTTGAAGAAATTA 580
QY 141 GlyHisThrSerGlyIleLysLysAlaLeuThrPysSerAspAspLysGlnIleLeuSer 160
DB 581 GGTATATCTGCTGATTAATAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
QY 161 AlaAspAspLysThrValArgLeuThrPaspHisAlaThrMetThrGluValLysSerLeu 180
DB 641 GCTGATGCCAAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
QY 181 AsnPheAsnMetSerValSerSerMetGluTyrIleProGluGlyGluIleLeuValIle 200
DB 701 AATTTTATATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 760
QY 201 ThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe 220
DB 761 ACTTATGAGCACTTATATGCTTNTCATAGTGCAGTAAGTTTGACCAATTAATCCTTT 820
QY 221 GluAlaProAlaThrIleAsnSerAlaSer**HisProGlu-LysGluPheLeuValAl 240
DB 821 TGAGCTCCGCAACCAATCATCTCATCTCTCATCTCTCATCTCTCATCTCTCTCT 880
QY 240 ArgLysGlyLysAspPheLysLeuTyr-LysTyrAspTyrIleAsnSerGly-LysGluL 258
DB 881 AGGCGCGGTGAATAATTTTAAACTTTTAAATAGTATGATTAATTAATGAGGAGGAAGAA 940
QY 258 eu-GluSerTyrLysGlyHis-LysPheGlyProIleHisCysValAlaPheSerProAsp 276
DB 941 TAAATATCTTACCAAGGAGAACACTTTGGGGGCC-----TAATTCACCGGGGG 988
QY 277 GlyGluLeuTyrAlaSerGlySerGlu 285
DB 989 GGAGAAATTAAGTCCCGGAGATGGAG 1015

```

RESULT 12
 AUI23939

```

LOCUS AUI23939 800 bp mRNA linear EST 01-AUG-2002
DEFINITION AUI23939 NT2RM2 Homo sapiens cDNA clone NT2RM2001349 5', mRNA
ACCESSION AUI23939
VERSION AUI23939.1 GI:10948655
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Salto,K., Yamamoto,J.,
1 (bases 1 to 800)
Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Salto,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Isogai,T.
HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Salto,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
,Y., Sugano,S., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomcs@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
SOURCE
1..800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM2001349"
/clone_1id="NT2RM2"
/cell_type="teratocarcinoma"
/note="Vector: pME18SF3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 227 a 168 c 194 g 211 t
ORIGIN
Alignment Scores: 5.89e-141 Length: 800
Pred. No.: 1203.00 Matches: 233
Score: 98.33% Conservative: 2
Percent Similarity: 97.49% Mismatches: 4
Best Local Similarity: 65.38% Indels: 1
Query Match: 9 Gaps: 0
DB:
US-09-856-836-2 (1-351) x AUI23939 (1-800)
QY 1 MetAlaMetArgGlnThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
DB 80 ATGGCAATGAGACAGACGCGCTACCTGCTGCGCACACGCCAGCCGCGTGGATTGG 139
QY 21 AlaPheSerGlyLeuThrProGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 140 GCCTTGAGTGCATCAGCGCTTATGGGTATTTCTTAATCAGCGCTGCAAGATGTAA 199
QY 41 PrometLeuArgGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
DB 200 CCTATGCTACGCGCAGGAGATACAGAGACTGATGGAACATTTTGGGTCAATAAGGT 259
QY 61 AlaValTrpGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaAlaAlaAsp 80
DB 260 GCTGTTGGGGGTGACACACTGATTAAGGATGCCAACAAAGCAGTACAGCGTGCAGAT 319
QY 81 PheThrAlaLysValTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 320 TTCACACCCAAAGTGGGATGCTGCTCAGAGATGAATGATGACCTGCGCTCATTA 379
QY 101 HisIleValLysThrValAspPheThrGlnAspSerAsnTyrLeu**ThrGlyGlyGln 120

```

```

Db 380 CACATTCCTCAACACTGCTGATTTTCACGACGAGATGATTAATTTGTTAAACCGGGGACAG 439
QY 121 AApLySLeuLeuArgIleTyraSpleuAnLysProGluAlaGluProLysGluLeuSer 140
Db 440 GATTAACCTGTCAGCATATATGACTTGAACAACCTGAAGCAACCTAAGAAATTTAGT 499
QY 141 GYHSTHrSerGlyIleLysLysAlaLeuTyPcYSerAspAspLysGluIleLeuSer 160
Db 500 GGTCTACTCTGCTATATAAAAAACCTCTGTGTGTCAGTGAAGATTAACAGATTTCTTCT 559
QY 161 AlaAspAspLysThrValArgLeuTrpAspHisAlaTrpMetThrGluValLysSerLeu 180
Db 560 GCTGATGACAAACACTGTTGCACTTTGGCATCATGCTACTATGACAGAAAGTAAATCTCTA 619
QY 181 AsnPheAsnMetSerValSerSerMetGluTyrlleProGluGluIleLeuValIle 200
Db 620 AATTTTAAATGCTGCTAGTAGTAGATGATATATCTCTGAGGGAGACATTTGCTTATA 679
QY 201 ThrTyrglyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe 220
Db 680 ACTTATGACCATCTATTTGCTTTTCATATGTCAGTAACTTGAGACCAATTAATCCCTTT 739
QY 221 GluAlaProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuVal 239
Db 740 GAAGCTCCTGC-ACCATCAATTCCTGATCTCTGATCTGAGAAAGAAATTTCTGTG 795

RESULT 13
Bg778828 796 bp mRNA linear EST 15-MAY-2001
LOCUS 602667790F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:480720 5'
ACCESSION mRNA sequence.
VERSION Bg778828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 796)
AUTHORS NIH-MGC http://mgs.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DPF
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LNCMI658 row: d column: 23
High quality sequence stop: 717.
Location/Qualifiers
1. 796
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="480720"
/clone_lib="NIH_MGC_60"
/tissue_type="adrenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: prostate; Vector: pDR-R-LIB (Clontech);
Site_1: SfiI (ggccgcctggcgc); Site_2: SfiI (ggccatagcgc
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCCATATGCGC-3' and 3' adaptor
sequence: 5'-ATCTTAGAGCGCGGCGCGCATG-dT(30)BN-3'
(Where B = A, C, or G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH-MGC

```

```

BASE COUNT 245 a 124 c 188 g 239 t
ORIGIN
Library: "
Alignment Scores:
Pred. No.: 1,07e-139 Length: 796
Score: 1193.00 Matches: 232
Percent Similarity: 97.12% Conservative: 4
Best Local Similarity: 95.47% Mismatches: 4
Query Match: 64.84% Indels: 3
DB: 12 Gaps: 0
US-09-856-836-2 (1-351) x Bg778828 (1-796)
QY 90 SerGlyAspGluMetThrLeuAlaHisLysHisIleValLysThrValAspPheTrp 109
Db 2 TCAAGATGATGAATGATGACCTGGCTGCATTAACATTTGTCAGACTGATTCACG 61
QY 110 GluAspSerAsnTyrlleu***ThrGlyGluAspLysLeuArgIleTyraSpleu 129
Db 62 CAGGATGATATTTATTTGTTAACCGGGGACAGATTAACCTGTACGATATATGACTTG 121
QY 130 AsnLysProGluAlaGluProLysGluIleSerGlyHisThrSerGlyIleLysLysAla 149
Db 122 AACAAACCTGAACAGACAACTAAGAAATTAAGTGCATATCTCTGTATAAAAAAGCT 181
QY 150 LeuTrpCysSerAspAspLysGluIleLeuSerAlaAspAspLysThrValArgLeuTrp 169
Db 182 CTGTGTGTCAGTAGAGATGAATTAACAGATCTTCTGTCAGAGACAAACCTGTGACTTTGG 241
QY 170 AspHisAlaThrMetThrGluValLysSerLeuAsnMetSerValSerSerMet 189
Db 242 GATCATGCTACTATGACAGAGAAATGTAATCTTAATTTAATATCTGTAGTACTATG 301
QY 190 GluTyrlleProGluGluIleLeuValIleThrTyrglyArgSerIleAlaPheHis 209
Db 302 GAATTAATTCCTGAGGAGAGATTTGTTAATACTTAGAGCATATTCCTTTTCAT 361
QY 210 SerAlaValSerLeuGluProIleLysSerPheGluAlaProAlaThrIleAsnSerAla 229
Db 362 AGTGCAGTAAAGTTTGACCAACCAATTAATCTTTGAAGCTCTGCAACCAATTCCTGCA 421
QY 230 Ser***HisProGluLysGluPheLeuValAlaGlyGlyLysAspPheLysLeuTyrls 249
Db 422 TCTCTTCATCTGAGAAATAATTTCTGTGAGGGGGGAGAAATTTAACTTTATAG 481
QY 250 TyrAspTyraSncSerGlyGluLeuGluSerTyrlsGlyHisPheGlyProIleHis 269
Db 482 TATGATTATTAATAGTGAGAGAAATTAAGATCTCAAGGACACTTGGTCTATTCAC 541
QY 270 CysValArgPheSerProAspGlyGluLeuTyraLysSerGlySer-GluAspGlyThrLe 289
Db 542 TGTGTGAGATTTAGTCCATATGAGACACTATGCAAGTGTTCAGAAATGAGAACATT 601
QY 289 uArgLeuTrpGluThrValIleGlyLysThrTyrlleu-LeuTyrlsCysVal***ProG 309
Db 602 GAGACTATGCGCAAACTGTGTGAGAAACGATAGCTTTGGAATAGTGCTCTCTG 661
QY 309 LysGluAspSer-GlyGluLeuAlaLysProLysIleGlyPheProGluThrAlaGluLys 328
Db 662 AAGAAGATATAGTGTGAGCTGCAACCAAGATTTGTTCCAGAGACAGAACAGAG 721
QY 329 Glu 329
Db 722 GAG 724

RESULT 14
LOCUS AL556515
DEFINITION AL556515 L1.NFL006_p12 Homo sapiens cDNA clone CSDBK006Yg21 5
prime, mRNA sequence.
ACCESSION AL556515
VERSION AL556515.1 GI:12899259
901 bp mRNA linear EST 16-FEB-2001

```


US-09-856-836-2 (1-351) x BQ433088 (1-1086)

```
QY      1 MetAlaMetArgGlnThrProLeuThrCysSerGlyHisThrArgProValValasp*** 20
      292 ATGGCATATGAGACAGACGCCGCTCAGCTGCTGCGCACACGCCGCGTGGTTGATTTG 351
QY      21 AlaPheSerGlyIleThrProTyrGlyIlePheLeuIleSerAlaCysLysAspGlyLys 40
      352 GCGTTCAGTGGGATCAGCGCTTATGAGTATTTCTTAATCAGCGCTTGCAGAAAGATGTAAA 411
QY      41 PrometLeuArgGlnGlyAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGly 60
      412 CCTATGCTACGCCAGGAGATACAGAGACTGATTCGACACTTTTGGCTCATAAAGGT 471
QY      61 AlaValIlePheGlyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
      472 GCTGTTGGGGTGCACACTGTAATAGGTGCCACCAAGCAGCTACAGACGCTGCAGAT 531
QY      81 PheThrAlaLysValIleTrpAspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
      532 TTCACAGCCAAAGTGGGATGCTCTCAGAGATGATGATGATGATGATGATGATGATGAT 591
QY      101 HisIleValLysThrValAspPheThrGlnAspSerAsnTyrLeu***ThrGlyGlyLys 120
      592 CACATTTGCAAGACTGTGATTTCCAGCAGATAGTAATTAATTTGTTAACCGGGGACAG 651
QY      121 AspLysLeuLeuArgGlyIleTyrAspLeuAsnLysProGluAlaGluProLysGluIleSer 140
      652 GATTAACCTGTTACGATTAATGACTGTAACCAACCTGTAACCAACCTGTAACCAACCT 711
QY      141 GlyHisThrSerGlyIleLysLysAlaLeuTyrPheSerAspLysGlnIleLeuSer 160
      712 GGTCACTACTTGGTATTAATAAAGCTCTGTGTGAGTGGAGATTAACAGATTTCTTCT 771
QY      161 AlaAspAspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeu 180
      772 GCTGATGACAAACTGTTCGACTTGGGATCATCTACTATGACAGCAAGTGAATCTCTA 831
QY      181 AsnPheAsnMetSerValSerSerMetGluTyrIleProGluGlyLysIleLeuValIle 200
      832 AATTTTAAATATGCTCTGTTAANAGNATGGATATATTCCTGAGGAGAGAT-TTGGTTATA 890
QY      201 ThrTyrGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPhe 220
      891 ACTTATGGACGATCTATGCTTTCT-AATGCCCTAGGTTTGGACCATTAATCCCTT 949
QY      221 GluAlaProAlaThrIleAsnSerAlaSer***HisProGlu-LysGluPheLeuValAl 240
      950 GAAAGTCTCCGAAACCAATCTGATCCCTCATCTGAAAGGAGGAGGAGGAGGAGGAGG 1009
QY      240 aglyGlyLysAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyLys-GluLeuGln 260
      1010 ACGGGGGGAAAATTTTAACCTTATTAAGTTTATTAATAAAGGGGAGGAGGAGGAGGAG 1069
QY      260 eTyrLysGlyHisPhe 265
      1070 TCTACAGGAGCACTTT 1086
Db
```

Search completed: February 7, 2003, 00:00:04
Job time : 2182 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2003, 13:03:57; Search time 57 Seconds

(without alignments)
1888.483 Million cell updates/sec

Title: US-09-856-836-2

Perfect score: 1840

Sequence: 1 MAMQPTLTCGHRFPVYDX.....EELASNSDSITSSTPEVKA 351

Scoring table:

BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USFTO.spool/US09856836/runat.31012003_144237_22183/app.query.fasta_1.519
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=tbl -MINMATCH=0.1 -LOOPLC=0
-LISTEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=tbl -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09856836 @CGL 1.1 17 @unat.31012003_144237_22183 -NCP=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGESUBSTR -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMECUT=120
-WARN_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA:*

1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
6: /cgn2_6/prodata/1/ina/5B.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1779.5	96.7	1820	2	US-08-828-922-2
2	425	23.1	686	4	US-09-040-984-36
3	425	23.1	686	4	US-09-123-912-36
4	425	23.1	686	4	US-09-643-597-36
5	279	15.2	5152	4	US-09-690-364-10
6	274	14.9	3747	4	US-09-690-364-17
7	268.5	14.6	7042	4	US-09-435-508-1
8	268.5	14.6	7042	4	US-09-435-508-1
9	268.5	14.6	7042	4	US-09-098-310-1
10	268.5	14.6	7042	4	US-09-690-364-21
11	268.5	14.6	7075	4	US-09-092-508-15
12	268.5	14.6	7075	4	US-09-435-115-15

13	238.5	13.0	2152	1	US-08-188-582-17	Sequence 17, Appl
14	238.5	13.0	2152	1	US-08-646-715-17	Sequence 17, Appl
15	221.5	12.0	2085	2	US-08-283-917-8	Sequence 8, Appl
16	221.5	12.0	2085	2	US-08-961-716-8	Sequence 8, Appl
17	200	10.9	2272	4	US-09-108-857-1	Sequence 1, Appl
18	195.5	10.6	1211	4	US-09-063-743-2	Sequence 2, Appl
19	195.5	10.6	1211	4	US-09-590-540-2	Sequence 2, Appl
20	192	10.4	2369	4	US-09-302-769-20	Sequence 20, Appl
21	191	10.4	1115	1	US-08-190-802A-19	Sequence 19, Appl
22	191	10.4	1115	1	US-08-473-346-19	Sequence 19, Appl
23	191	10.4	1115	4	US-08-473-346-19	Sequence 19, Appl
24	187	10.2	2186	4	US-08-487-072A-19	Sequence 19, Appl
25	187	10.2	2558	4	US-09-184-001-1	Sequence 19, Appl
26	187	10.2	2558	4	US-09-184-001-1	Sequence 19, Appl
27	187	10.2	7881	2	US-08-751-189-1	Sequence 1, Appl
28	187	10.2	7881	2	US-09-060-836-1	Sequence 1, Appl
29	187	10.2	7881	4	US-09-184-445-1	Sequence 1, Appl
30	185.5	10.1	1766	5	PCT-US93-00601-1	Sequence 1, Appl
31	185.5	10.1	1766	5	PCT-US94-07107A-1	Sequence 1, Appl
32	185.5	10.1	1767	4	US-08-083-945C-1	Sequence 1, Appl
33	185	10.1	1246	4	US-09-302-769-22	Sequence 2, Appl
34	185	10.1	7886	2	US-08-751-189-2	Sequence 2, Appl
35	185	10.1	7886	2	US-09-060-836-2	Sequence 2, Appl
36	185	10.1	7886	2	US-09-184-445-2	Sequence 2, Appl
37	182.5	9.9	1517	4	US-09-180-783-1	Sequence 1, Appl
38	182	9.9	1394	4	US-09-147-826B-1	Sequence 1, Appl
39	182	9.9	2119	3	US-09-032-372-6	Sequence 6, Appl
40	179.5	9.8	1368	3	US-08-707-339E-1	Sequence 1, Appl
41	179	9.7	2359	1	US-08-188-582-4	Sequence 1, Appl
42	179	9.7	2359	1	US-08-646-715-4	Sequence 4, Appl
43	179	9.7	2481	3	US-08-899-578-1	Sequence 1, Appl
44	175.5	9.5	1020	3	US-08-707-339E-3	Sequence 1, Appl
45	174	9.5	2369	2	US-08-883-534-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-828-922-2
Sequence 2, Application US/08828922
Patent No. 5834240

GENERAL INFORMATION:
APPLICANT: Olga, Bandman
TITLE OF INVENTION: TRANSFORMING GROWTH FACTOR-B RECEPTOR
TITLE OF INVENTION: ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,922
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0258 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

DB	Sequence	Length	Matches	Conservative	Mismatches	Indels	Gaps
US-08-828-922-2	1 MetAlaMeFArGlnThrProLeuThrCysSerGlyHisThrArgProValAlaAsp***	1,212-209	1820	337			
US-08-828-922-2	2						
US-08-828-922-2	3						
US-08-828-922-2	4						
US-08-828-922-2	5						
US-08-828-922-2	6						
US-08-828-922-2	7						
US-08-828-922-2	8						
US-08-828-922-2	9						
US-08-828-922-2	10						
US-08-828-922-2	11						
US-08-828-922-2	12						
US-08-828-922-2	13						
US-08-828-922-2	14						
US-08-828-922-2	15						
US-08-828-922-2	16						
US-08-828-922-2	17						
US-08-828-922-2	18						
US-08-828-922-2	19						
US-08-828-922-2	20						
US-08-828-922-2	21						
US-08-828-922-2	22						
US-08-828-922-2	23						
US-08-828-922-2	24						
US-08-828-922-2	25						
US-08-828-922-2	26						
US-08-828-922-2	27						
US-08-828-922-2	28						
US-08-828-922-2	29						
US-08-828-922-2	30						
US-08-828-922-2	31						
US-08-828-922-2	32						
US-08-828-922-2	33						
US-08-828-922-2	34						
US-08-828-922-2	35						
US-08-828-922-2	36						
US-08-828-922-2	37						
US-08-828-922-2	38						
US-08-828-922-2	39						
US-08-828-922-2	40						
US-08-828-922-2	41						
US-08-828-922-2	42						
US-08-828-922-2	43						
US-08-828-922-2	44						
US-08-828-922-2	45						
US-08-828-922-2	46						
US-08-828-922-2	47						
US-08-828-922-2	48						
US-08-828-922-2	49						
US-08-828-922-2	50						
US-08-828-922-2	51						
US-08-828-922-2	52						
US-08-828-922-2	53						
US-08-828-922-2	54						
US-08-828-922-2	55						
US-08-828-922-2	56						
US-08-828-922-2	57						
US-08-828-922-2	58						
US-08-828-922-2	59						
US-08-828-922-2	60						
US-08-828-922-2	61						
US-08							

Qy	261	TYLSTSLYLSHSHGELIPROLEHSHSCYSVALATGPHSESRPROASGIVGLEUENYR	280
Db	1039	TACAGGGAACACTTGGCTCTATTCACCTGTGACATTATAGCCGAGGAGAACTCAT	1098
Qy	281	ALASERGLYSERGIUASPGLYTHRLEUATGLEUTPGINTHRVALVAGLYLSTHRYR	300
Db	1099	GCCAGTGCTTCAGACAGATGGACATCTGAGCAAACTGTGTGAGAAAAACGTAT	1158
Qy	301	GLYEURTRPYSCYSVAL**PROGLUGIUAASERGLGLEUVALALYSERPROLYSILE	320
Db	1159	GGCCTTTGGAAATGTGTGCTTCTGTAAAGAAAGATAGTGTGACTGTGGCAAGCCAAAGATT	1218
Qy	321	GLYPHEPROGLUTHRALAGIUGIUGIULEUVALGILUGIILEALASERGIUASNSERSP	340
Db	1219	GGTTCCTCCAGAGAACAGGAGAGAGACTA--GAAGAAATATGCTTCACAGAAATTCAGAT	1275
Qy	341	SERLETRYSERSETRHRRPGIUGIUALYSALA	351
Db	1276	TGCACCTTTCCTTCAGCTCCTGATGATTAAAGCC	1308

RESULT 2

```

US-09-040-984-36
: Sequence 36, Application US/09040984
: Patent No. 6210883
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Wang, Tongtong
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSTICS
: TITLE OF INVENTION: OF LUNG CANCER
: NUMBER OF SEQUENCES: 86
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED AND BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSO for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/040,984
: FILING DATE: 18-MAR-1998
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Maki, David J.
: REGISTRATION NUMBER: 31,392
: REFERENCE/DOCKET NUMBER: 210121.456
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-622-4900
: TELEFAX: 206-282-6031
: TELEX:
: INFORMATION FOR SEQ ID NO: 36:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 686 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
US-09-040-984-36
:
: Alignment Scores:
: Pred. No.: 4.04e-43 Length: 686
: Score: 425.00 Matches: 107
: Percent Similarity: 83.97% Conservative: 3
: Best Local Similarity: 81.68% Mismatches: 21
: Query Match: 23.10% Indels: 7
: Gaps: 0
:
US-09-856-836-2 (1-351) x US-09-040-984-36 (1-686)

```

```
OY 1 MetAlaMetArgGlnThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
DB 277 ATGGCAATNAGACAGACCGCGCTGCTCTGGGCGACACCGCGCGCGTGGTATTG 336
OY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 337 GCCTTCAGTGGCATCAC-CCTTATGGGTATTCTTATCATCAGCGCTTGCAGAGATGGTTAA 395
OY 41 PrometLeuArgGlnGlyAspThrGlyAspTyrIleGlyThrPheLeuGlyHisLysGly 60
DB 396 CCTATGCTACGCGCAGGAGATACAGACACTGATGGAACATTTTGGGGCTCTAAAGCT 455
OY 61 AlaValTyrPglyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
DB 456 CT-GTTTGGGGTGCAACCTGATGATAGATGCGCACCAAGCAGCTACACAGCTTCAGAT 514
OY 81 PheThrAlaLysValTyrPaspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 515 TTCACAGCCCAAGTGTGGATGCTCTCAGAGANTATGATTAACCTG-GCTCATTA- 572
OY 101 HisIleValLysThrValAspPheThrGlnAspSerAsnTyrLeu***ThrGlyGlyGln 120
DB 573 CACATTTGCAAGATGTGATTTCCC-CAGGATAT-TATTATTGTGTACCGGGGAGANAG 630
OY 121 AspLysLeuLeuArgIleTyrAspLeuAsnLys 131
DB 631 GATTA-CTGTTTCNCNTATTATTAATGAACAA 662

RESULT 3
US-09-123-912-36
: Sequence 36, Application US/09123912A
: Patent No. 6312695
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
: FILE REFERENCE: 210121.455C1
: CURRENT APPLICATION NUMBER: US/09/123,912A
: CURRENT FILING DATE: 1998-07-27
: PRIOR APPLICATION NUMBER: 09/040,802
: PRIOR FILING DATE: 1998-03-18
: NUMBER OF SEQ ID NOS: 114
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 36
: LENGTH: 686
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (222)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (224)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (237)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (264)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (285)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (548)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (551)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (628)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
```

```
: LOCATION: (643)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (645)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (665)
: OTHER INFORMATION: where n is a, c, g or t
: NAME/KEY: modified_base
: LOCATION: (674)
: OTHER INFORMATION: where n is a, c, g or t
US-09-123-912-36

Alignment Scores:
Pred. No.: 4,04e-43 Length: 686
Score: 425.00 Matches: 107
Percent Similarity: 83.97% Conservative: 3
Best Local Similarity: 81.68% Mismatches: 21
Query Match: 23.10% Indels: 7
DB: 4 Gaps: 0

US-09-856-836-2 (1-351) x US-09-123-912-36 (1-686)

OY 1 MetAlaMetArgGlnThrProLeuThrCysSerGlyHisThrArgProValAlaAsp*** 20
DB 277 ATGGCAATNAGACAGACCGCGCTGCTCTGGGCGACACCGCGCGTGGTATTG 336
OY 21 AlaPheSerGlyIleThrProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLys 40
DB 337 GCCTTCAGTGGCATCAC-CCTTATGGGTATTCTTATCATCAGCGCTTGCAGAGATGGTTAA 395
OY 41 PrometLeuArgGlnGlyAspThrGlyAspTyrIleGlyThrPheLeuGlyHisLysGly 60
DB 396 CCTATGCTACGCGCAGGAGATACAGACACTGATGGAACATTTTGGGGCTCTAAAGCT 455
OY 61 AlaValTyrPglyAlaThrLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAsp 80
DB 456 CT-GTTTGGGGTGCAACCTGATGATAGATGCGCACCAAGCAGCTACACAGCTTCAGAT 514
OY 81 PheThrAlaLysValTyrPaspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 515 TTCACAGCCCAAGTGTGGATGCTCTCAGAGANTATGATTAACCTG-GCTCATTA- 572
OY 101 HisIleValLysThrValAspPheThrGlnAspSerAsnTyrLeu***ThrGlyGlyGln 120
DB 573 CACATTTGCAAGATGTGATTTCCC-CAGGATAT-TATTATTGTGTACCGGGGAGANAG 630
OY 121 AspLysLeuLeuArgIleTyrAspLeuAsnLys 131
DB 631 GATTA-CTGTTTCNCNTATTATTAATGAACAA 662

RESULT 4
US-09-643-597-36
: Sequence 36, Application US/09643597
: Patent No. 6426072
: GENERAL INFORMATION:
: APPLICANT: Wang, Tonglong
: APPLICANT: Fan, Liqun
: APPLICANT: Kalos, Michael D.
: APPLICANT: Bangur, Chaitanya S.
: APPLICANT: Hosken, Nancy
: APPLICANT: Fanger, Gary R.
: APPLICANT: Li, Samuel X.
: APPLICANT: Wang, Aijun
: APPLICANT: Skeiky, Yasir A.W.
: APPLICANT: Henderson, Robert A.
: APPLICANT: McNeill, Patricia D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.455C1
: CURRENT APPLICATION NUMBER: US/09/643,597
: CURRENT FILING DATE: 2000-08-21
: NUMBER OF SEQ ID NOS: 369
```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 686
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)---(686)
; OTHER INFORMATION: n = A,T,C or G
US-09-643-597-36

Alignment Scores:
Pred. No.:      4,04e-43          Length:      686
Score:           425.00          Matches:     107
Percent Similarity: 83.97%       Conservative: 3
Best Local Similarity: 81.68%    Mismatches: 21
Query Match:      23.10%         Indels:      7
DB:               4              Gaps:          0

US-09-856-836-2 (1-351) x US-09-643-597-36 (1-686)

QY   1 MetAlaMetArGInThrProLeuThrcysSerGIyHisThrArProValIAsp*** 20
      |||||
Db   277 ANGCAGATNAGCAACAGCCCGCTCACCTGCTGTGGCAGACGCCGCGTGTATTG 336
      |||||

QY   21 AlaPheSerGIyIleThrProTyGIyTyPhelLeuIleSerAlaCyLySAspGIyLys 40
      |||||
Db   337 GCCTTCAGTGCGATCAC-CCTTATGGGTAATTTCTTAATCACCGCTTCAAAGATGGTTAA 395
      |||||

QY   41 PrometLeuArGInGIyAspThrGIyAspTrpIleGIyThrPhelLeuGIyHisLySGly 60
      |||||
Db   396 CCTATGCTACGCCAAGCGAGATACAGAGACTGATGTGAACATTTTTGGGGGTCTAAAGGT 455
      |||||

QY   61 AlAlaValTPrgIyAlaThrlLeuAsnLysAspAlaThrlLysAlaIalThrAlaIalaAsp 80
      |||||
Db   456 CT-GTTTGGGGTGCACAACCTGATTAAGATGCCACAAAGCAGCTACAGCAGTGCAGAT 514
      |||||

QY   81 PheThrAlaLysValITrpAspAlaValASerGIyAspLeuMetThrLeuAlaHisLys 100
      |||||
Db   515 TTCACAGGCCAACAGTGGGATGCTGTCCAGANATTAATTGATAACCTG-CCTCATTA- 572
      |||||

QY   101 HisIleValLysThrValAspPheThrcInGlnAspSerAsnTyrlLeu***ThrlGIyGln 120
      |||||
Db   573 CACATGTTCACAGAAATGTGATTTCCC-CAGGATAT-TATTATTGTTTACC GGGGGANAG 630
      |||||

QY   121 AspLysLeuLeuArgIleTyrrAspLeuAsnLys 131
      ||| |||
Db   631 GATTA-CTGTTNCNTATTTTAATTGAACAA 662
      ||| |||

RESULT 5
US-09-690-364-10
; Sequence 10, Application US/09690364
; Patent No. 6468795
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; APPLICANT: Andrew T. Walt
; TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
; FILE REFERENCE: RTS-0190
; CURRENT APPLICATION NUMBER: US/09/690,364
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 100
; SEQ ID NO 10
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (886)...(4302)
US-09-690-364-10

Alignment Scores:
Pred. No.:      1.08e-23          Length:     5152
Score:           279.00          Matches:     83
```

Percent Similarity:	44.2%	Conservative:	51
Best Local Similarity:	27.39%	Mismatches:	143
Query Match:	15.16%	Indels:	26
DB:	4	Gaps:	8
US-09-856-836-2 (1-351) x US-09-690-364-10 (1-5152)			
QY	13 HisHrAngrProValValAsp***AlaPheSerGlyIleThrProTyrGlyTyrPheLeu	32	
DB	2392 CACACAGATGCTGTTTACACACGCGCTGTTTTCTT-----CAGATGGCCAGAGAA	2442	
QY	33 IleSerAlaCysLysAspGlyLysPrometLeuArgGlnGlyAspThrGlyAspTrpIle	52	
DB	2443 GCTTCTTGCGGGCGGTATTAACCTTAACAGTGTTAAGACCCGAGACAGAGAAACTT	2502	
QY	53 GlyThrPheLeuGlnLysLysGlyAlaValTyrGlyAlaThrLeuAsnLysAspAlaThr	72	
DB	2503 CTTCACATTAAAGCTCATGAAGATGAGGTCTCTGCGCGCTCTCCACAGACAGACT	2562	
QY	73 LysAlaAlaThrAlaAlaAlaAspPheThrAlaLysValTyrAspAlaValSerGlyAsp	92	
DB	2563 TACATAGAGACCTCTCCAGCGGATGAAGACGTAAAGTTTGGATTTCCGACCTGGAG	2622	
QY	93 GluLeuMetThrLeu---AlaHisLysHisIleValLysThrValAspPheThrGlnAsp	111	
DB	2623 CTTCGACACACTTCACGACGACACTCGGACCAAGTAATGCTGCCATTTCACCAACGT	2682	
QY	112 SerAsnTyr-----Leu***ThrGlyGlnAspLysLeuLeuArgIleTyrAspLeu	129	
DB	2683 AGTAAACACCTCTCTCTGGGACCGGGTCAAAAGATTCTTCTCTCAGCTCGGATTTG	2742	
QY	130 AsnLysProGluAlaGluProLysGluLysSerGlyHisThrSerGlyIleLysLysAla	149	
DB	2743 AATCAAAAGAAATGCA---AATACCACTGTTGGTGCACACCACTACGACCACTGC	2799	
QY	150 LeuTrpCysSerAspAspLysGlnIleLeuSerAlaAsp---AspLysThrValArgLeu	168	
DB	2800 AGGTTCTACCAACGACGATGAGCTCTTGCGTCTAGTGTCTGACGAGGACCTTAAGGCTT	2859	
QY	169 TrpAspHisAlaThrMetThrGluValLysSerLeuAsnPheAsnMetSerValSerSer	188	
DB	2860 TGGGATTCGATCATCGCAAAACGAGAGAAAGACATTAAATGTAAGCGCTTCTCTGAGT	2919	
QY	189 MetGluTyrIleProGlu-----	194	
DB	2920 TCACAGAACCTCTCCAGAGAGATGTGAGGTGATCGTGAAGTGTGTCCTCGCTGCAGAT	2979	
QY	195 ---GlyGluIleLeuValIleThrTyrGlyArgSerIleAlaPhe-----HisSerAla	211	
DB	2980 GGTGACAAATTAATATGAGGACGACGCAAAACAAAGTCTCTTTTGATTTATATCTAGT	3039	
QY	212 ValSerLeuGluProIleLysSerPheGluAlaProAlaThrIleAsnSerAlaSer***	231	
DB	3040 GGCCTATTGGCAGAGATCCACACAGGCAAT---CACAGACCATTCAGACTGATGATCTC	3086	
QY	232 HisProGluLysGluPheLeuValAlaGlyGlyLysAspPheLysLeuTyrLysTyrAsp	251	
DB	3097 TCCCCCTATGACCAATTGGCTGTGATGGCTCTCTCAGTACGTGTGAGAGTTGCGAAC	3156	
QY	252 TyrAsnSerGlyGluLeuGluLeuGluSerTyrLysGlyHisPheGlyProIleHisCysVal	271	
DB	3157 ATAGACTCCGCCCTAAAGGAGGCGCAGCTGACAGAGCAATTGAGTTGAGTTACACGCTG	3216	
QY	272 ArgPheSerProAspGlyGluLeuTyrAlaSerGlySerGluAspGlyThrIleuArgLeu	291	
DB	3217 ATGTTTCTCCGATGGCTCTCATTTTGCACGCTTGATGACCAACATATAGGCTC	3276	
QY	292 TrpGlnThr 294		
DB	3277 TGGGAGACA 3285		

Sequence 17, Application us/09690364

Patent No. 6468795

GENERAL INFORMATION:

APPLICANT: Hong Zhang

APPLICANT: Andrew T. Walt

TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION

FILE REFERENCE: RTS-0190

CURRENT APPLICATION NUMBER: US/09/690,364

CURRENT FILING DATE: 2000-10-17

NUMBER OF SEQ ID NOS: 100

SEQ ID NO 17

LENGTH: 3747

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(3747)

US-09-690-364-17

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-856-836-2 (1-351) x US-09-690-364-17 (1-3747)

7 ProleuthrCys-SerGlyHisThrArgProValAlaAsp**AlaPheSerGlyLeuTh

1821 CCGGTAGTGTGCGCGCCACACAGATGCTTTACCAAGCGCTTTCT-----1872

26 rProtyrGlyThrPheLeuLeuSerAlaCysLysAspGlyLysPheLeuArgGlnG1

1873 -GAGGATGTCAGAGATGATCTTCTTGAGAGCTGATAAACCTTACAGGCTTCAAGC

46 yAspThrGlyAspThrLeuGlyThrPheLeuGlyHisLysGlyAlaValThrPglYalTh

1932 TGAACGAGAGAGAACTCTAGCAAGGCTCAGAGATGAGAGTCTTGTGTGTC

66 rLeuAnLysAspAlaThrLysAlaAlaThrAlaAlaAlaAspPheThrAlaLysValTr

1992 ATCTCTACAGATGAGATTTATACCAACTCTCAGTGGATGATAAAGTGAAGATTG

86 pAspAlaValSerGlyAspGluLeuMetThrLeu---AlaHisLysHisLeuValLysTh

2052 GAATTCATGACTGGGAGACTAGTACACACTTGTGAGCACTCAGACCAAGTCAATTG

105 rValAspPheThrGlnAspSerAsnTr-----Leu**ThrGlyGlyLysAspLysLe

2112 CTGCCATTTCACCAACAGATGATCATCTTCTTCAAGCAGTGGTCAAGTGCCTT

123 uLeuArgLysLeuAspLeuAnLysProGluAlaGluProlYsgLuiLysSerGlyHisTh

2172 CCTCAAACTTTGGGATTGAATCAAAAAGAAAGTGA--AATACACTTTTGGTCATAC

143 rSerGlyLeuLysLysAlaLeuThrPcySerAspLysGlnLeuSerAlaAsp--162

2229 AATTCAGTCATCACTGAGATTTTCACCAATGATTAAGCTTTGGCTAGTGTTCAGC

163 -AspLysThrValArgLeuThrPAspHisAlaThrMetThrGluValLysSerLeuAnPh

2289 TGAATGAACTTAAAGCTTTGGGATGCGACATCAAAATGAGAGAAAGACATTAAAGT

182 eAsnMetSerValSerSerMetGluTrLLeuProGlu-----194

2349 GAAACAGTCTTCCTTAATTTGGAGAGACCTTCAAGAGATATGAGAGTATGAGTGAAGTG

195 -----GlycLuiLeuValLLeuThrTrGlyArgSerLLeuAl 207

2409 TTGTTCTGTGTCTGCTGATGATGCAAGATATATGCTGCAGCAAAAATTAATCTTTCT 2468

QY 207 aphe-----HisSerAlaValSerLeuGluProLleLysSerPheGluAlaProAlaTh 225

2469 TTTTGACATTCATACATGTCGGCTATTGGAGAAATCCACACCGGCCAT---CACAGCAC 2525

225 rLeuAnSerAlaSer**HisProGluLysGluPheLeuAlaLysGlyLysAspPh 245

2526 CATTCCAGTACTGATCTTCTCCCAAAAACATTGGCAGTGTGCTTTGGCCAGTA 2585

245 eLysLeuTrLysTrAspPyrAsnSerGlyGluGluLeuGluSerTrLysGlyHisPh 265

2586 CTGTGTAGACTTGTGGAATACAGATCAGCTTCAAGGTCGATGAGAGACATTTT 2645

265 eGlyProLleHisCysValArgPheSerProAspGlyGluLeuTrAlaSerGlySerG1 285

2646 AAGTTGGTTCATGATGATGATGATTTCTCTCTGATGATCATCATTTTACATCTTCTGA 2705

285 uAspGlyThrLeuArgLeuThrProLlnThr 294

2706 TGACACAGACATCATGAGCTCTGGAGACA 2733

RESULT 7

US-09-092-508-1

Sequence 1, Application us/09092508

Patent No. 6291643

GENERAL INFORMATION:

APPLICANT: Henzel, William J.

TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt

STREET: 3100 No. 6291643 West Center, 90 South Seventh St

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/092,508

FILING DATE: 05-JUN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/048,807

FILING DATE: 05-JUN-1997

APPLICATION NUMBER: 60/055,258

FILING DATE: 07-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: Ketelberger, Ph.D., Denise M

REGISTRATION NUMBER: 33,924

REFERENCE/DOCKET NUMBER: 11669,6USU1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-332-5300

TELEFAX: 612-332-9081

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 7042 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 578...4159

OTHER INFORMATION:

US-09-092-508-1

Alignment Scores:

Pred. No.:

Score:

3,51e-22

Length:

7042

Matches:

77

Percent Similarity: 44.03% Conservative: 52
 Best Local Similarity: 26.28% Mismatches: 128
 Query Match: 14.59% Indels: 36
 DB: 4 Gaps: 7

US-09-856-836-2 (1-351) x US-09-092-508-1 (1-7042)

QY 7 ProLeuThrcys-SerGlyHisThrArgProValValAsp***AlaPheSerGlyIleTh 26
 Db 2365 CCGCTTAGTGTGCGCCGCCACACAGATGCGTTTACCAGCTGCTTTCT----- 2416
 QY 26 rProTyGlyTyrrPheLeuIleSerAlaCysLysAspGlyLysPrometLeuArgIngl 46
 Db 2417 -GAGGATGGTCAGGAATAGCTTTGTGGAGCTGATAAACCTTACAGGTGTCAAAGC 2475
 QY 46 yAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGlyAlaValITrpglyAlaTh 66
 Db 2476 TGAACACAGGAGAAACTTTCAGAAATCAAGCTCATGAGCATGAAGTCTTTGTGTGC 2535
 QY 66 rLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTr 86
 Db 2536 ATTCTCTACAGATACAGATTTATAGCACTGCTGCTAGTAAATAAAGTGAAGATTG 2595
 QY 86 pAspAlaValSerGlyAspGlyLeuMetThrLeu---AlaHisLysHisIleValLysTh 105
 Db 2596 GAATTCATGACCTGGGAGAACTAGACACACCTATGATGAGCTCAGACAGCAAGTCAATTG 2655
 QY 105 rValAspPheThrGlnAspSerAsnTy-----Leu**ThrGlyGlyLysAspLysLe 123
 Db 2656 CTGCCATTTTCACCACTAGCTCATCATCTCTCTTACCGCTGCTGCTGCTGCTT 2715
 QY 123 uLeuArgTleLysAspLysLeuAsnLysProGluAlaGluProLysGluIleSerGlyHisTh 143
 Db 2716 CCTCAACTTTGGGATTTGATCAAAAAGAAATGTGA---AATACCAGTTTGGTCATAC 2772
 QY 143 rSerGlyIleLysLysAlaLeuTrpCysSerAspLysGlnIleLeuSerAlaAsp-- 162
 Db 2773 AAATTCACCTCATCTCAGCTCAGATTTTCCACAGATGATTAAGCTTTGGCTAGTTCAGC 2832
 QY 163 -AspLysThrValArgLeuTrpAspHisAlaThrMetThrGlyValLysSerLeuAspTh 182
 Db 2833 TGAATGGAACCTTAAAGCTTTGGATGCGACATCAGCAAAATGAGAGAAAGCATTAATCT 2892
 QY 182 eAsnMetSerValSerSerMetGluTrpIleProGluGlyIleLeuValIleThrTy 202
 Db 2893 GAACAGTTCCTCCAAATTTGGAG-----GACCTCAAGAGCATATGGAAGT 2917
 QY 202 rGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPheGlnI 222
 Db 2918 -----GACCTCAAGAGCATATGGAAGT 2940
 QY 222 aProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuAlaAlaGlyL 242
 Db 2941 GATAGTGAAGTGTGTGCTGCTGCTGATGTCGACAGATTAATGCTGCCAGCAAAAA 3000
 QY 242 yGluAspPheLysLeuTyrrLysTyrrAspLysSerGlyGluGluLeuGlnSerTyrr 262
 Db 3001 T-----AAATCTTTTGTGGAAATACAGATCAGCTCAAGGCTGATGTCAG 3051
 QY 262 sGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyLysLeuTyrrAla 282
 Db 3052 AGGACATTTTAAGTGGTTCATGCTGTGATGTTTCTCCGATGATCATCATTTTTCAC 3111
 QY 282 rGlySerGlyAspGlyTrpLeuArgLeuTrpGlnTr 294
 Db 3112 ATCTTCTGATGACAGCAATCAGCTCTGGAGACA 3148

RESULT 8
 US-09-435-115-1
 : Sequence 1, Application US/09435115
 : Patent No. 6346607
 : GENERAL INFORMATION:
 : APPLICANT: Henzel, William J.

1 TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF CASPASE-3
 2 NUMBER OF SEQUENCES: 16
 3 CORRESPONDENCE ADDRESS:
 4 ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
 5 STREET: 3100 No. 634607 West Center, 90 South Seventh St
 6 CITY: Minneapolis
 7 STATE: MN
 8 COUNTRY: USA
 9 ZIP: 55402
 10 COMPUTER READABLE FORM:
 11 MEDIUM TYPE: Diskette
 12 COMPUTER: IBM Compatible
 13 OPERATING SYSTEM: DOS
 14 SOFTWARE: FastSeq for Windows Version 2.0
 15 CURRENT APPLICATION DATA:
 16 APPLICATION NUMBER: US/09/435,115
 17 FILING DATE:
 18 CLASSIFICATION:
 19 PRIORITY APPLICATION DATA:
 20 APPLICATION NUMBER: 09/092,508
 21 FILING DATE:
 22 APPLICATION NUMBER: 60/055,258
 23 FILING DATE: 07-AUG-1997
 24 ATTORNEY/AGENT INFORMATION:
 25 NAME: Kettelberger, Ph.D., Denise M
 26 REGISTRATION NUMBER: 33,924
 27 REFERENCE/DOCKET NUMBER: 11669,6USU1
 28 TELECOMMUNICATION INFORMATION:
 29 TELEPHONE: 612-332-5300
 30 TELEFAX: 612-332-9081
 31 TELEX:
 32 INFORMATION FOR SEQ ID NO: 1:
 33 SEQUENCE CHARACTERISTICS:
 34 LENGTH: 7042 base pairs
 35 TYPE: nucleic acid
 36 STRANDEDNESS: double
 37 TOPOLOGY: linear
 38 MOLECULE TYPE: CDNA
 39 FEATURE:
 40 NAME/KEY: Coding Sequence
 41 LOCATION: 578...4159
 42 OTHER INFORMATION:
 43 US-09-435-115-1
 44 Alignment Scores:
 45 Pred. No.: 3,51e-22 Length: 7042
 46 Score: 268.50 Matches: 77
 47 Percent Similarity: 44.03% Conservative: 52
 48 Best Local Similarity: 26.28% Mismatches: 128
 49 Query Match: 14.59% Indels: 36
 50 DB: 4 Gaps: 7
 51 US-09-856-836-2 (1-351) x US-09-435-115-1 (1-7042)
 52 QY 7 ProLeuThrcys-SerGlyHisThrArgProValValAsp***AlaPheSerGlyIleTh 26
 53 Db 2365 CCGCTTAGTGTGCGCCGCCACACAGATGCGTTTACCAGCTGCTTTCT----- 2416
 54 QY 26 rProTyGlyTyrrPheLeuIleSerAlaCysLysAspGlyLysPrometLeuArgIngl 46
 55 Db 2417 -GAGGATGGTCAGGAATAGCTTTGTGGAGCTGATAAACCTTACAGGTGTCAAAGC 2475
 56 QY 46 yAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGlyAlaValITrpglyAlaTh 66
 57 Db 2476 TGAACACAGGAGAAACTTTCAGAAATCAAGCTCATGAGCATGAAGTCTTTGTGTGC 2535
 58 QY 66 rLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTr 86
 59 Db 2536 ATTCTCTACAGATACAGATTTATAGCACTGCTGCTAGTAAATAAAGTGAAGATTG 2595
 60 QY 86 pAspAlaValSerGlyAspGlyLeuMetThrLeu---AlaHisLysHisIleValLysTh 105
 61 Db 2596 GAATTCATGACCTGGGAGAACTAGACACACCTATGATGAGCTCAGACAGCAAGTCAATTG 2655


```

QY 105 rValaspheThrglnaspSerAsnTyr-----Leu***ThrGlyGlyGlnAspLysle 123
DB 2656 CTGCCATTTCACCAACAGATGATCATCTCTTACGACTGGGTCAAGGACTGCTT 2715
QY 123 uLeuArgIleTyraSpLeuAsnLysProGlnAgluProLysGluIleSerGlyHisTh 143
DB 2716 CCTCAAACTTTGGATTTGAATCAAAAGAAAGTGTGA--AAATACCATGTTTGTCAATC 2772
QY 143 rSerGlyIleLysLysAlaLeuTrpCysSerAspLysGlnIleLeuSerAlaAsp-- 162
DB 2773 AAATTCAGTCATCACTGCAATTTTCAACCATATTAAGCTTTGGCTAGTTGTTCC 2832
QY 163 -AspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeuAsnPh 182
DB 2833 TCATGGAACCTTAAAGCTTTGGATGCGACATCAACAATGAGAGCAAAAGCATTAATGT 2892
QY 182 eAsnMetSerValSerSerMetGluTyrIleProGluGluIleLeuValIleThrTy 202
DB 2893 GAAACAGTTCTCTTAATTTGGAG----- 2917
QY 202 rGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPheGluAl 222
DB 2918 -----GACCTCAAGAGCATATGGAAGT 2940
QY 222 aProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAlaGlyG 242
DB 2941 GATAGTGAAGTGTGTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 3000
QY 242 yGluAspPheLeuTyraSpLysTyraSpLysSerGlyLysGluLysGluLysSerTyry 262
DB 3001 T-----AAATCTTTTGTGGAATACAGACTCACGTTCAAAAGTGGCTGATTCGAC 3051
QY 262 sGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyLysGluLysTyry 282
DB 3052 AGGACATTTAAGTTGGTGGTGGATGATGATGATGATGATGATGATGATGATGATG 3111
QY 282 rGlySerGluAspGlyThrLeuArgLeuTrpGlnThr 294
DB 3112 ATCTTCTGATGACCAACATCAGGCTCTGGGAGACA 3148

RESULT 9
US-09-098-310-1
; Sequence 1, Application US/09098310
; Patent No. 6403765
; GENERAL INFORMATION:
; APPLICANT: Alnemt1, Emd S.
; TITLE OF INVENTION: TRUNCATED APAF-1 AND METHODS OF USE
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: 480140.438
; CURRENT APPLICATION NUMBER: US/09/098,310
; CURRENT FILING DATE: 1998-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7042
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (578)...(4159)
; US-09-098-310-1

Alignment Scores:
Pred. No.: 3, 51e-22 Length: 7042
Score: 268.50 Matches: 77
Percent Similarity: 44.038 Conservative: 77
Best Local Similarity: 26.288 Mismatches: 128
Query Match: 14.598 Indels: 36
DB: 4 Gaps: 7
US-09-856-836-2 (1-351) x US-09-098-310-1 (1-7042)

```

```

QY 7 ProLeuThrCys-SerGlyHisThrArgProValValAsp***AlaPheSerGlyIleTh 26
DB 2355 CCGCTTACTTTCCGCCCCCAACAGATGCTGTTTACACGCTGCTTCTCT----- 2416
QY 26 rProTyrlleTyraPheLeuIleSerAlaCysLysAspGlyLysProMetLeuArgGln 46
DB 2417 -GAGATGTCGACAGAAATAGCTTCTTGTGAGCTGATTAACCTTACAGGTTCCAAAGC 2475
QY 46 yAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGlyValAlaValTrpGlyAla 66
DB 2476 TGAACACGAGAGAACTTCTTCAAAATCAAGGCTTCAAGAGATGAATGATGCTTGTGTC 2535
QY 66 rLeuAsnLysAspAlaThrLysAlaIleAlaIleAlaAspPheThrAlaLysValTr 86
DB 2536 ATTCTTACAGATACAGATTTATAGCAACCTGCTCAGTGAATTAATAAAGTGAAGATTG 2595
QY 86 rAspAlaValSerGlyAspGluLeuMetThrLeu--AlaHisLysHisIleValLysTh 105
DB 2596 GAATTCATGACTGGGAGACTAGTACACACCTATGATGACACTCAGACGATCAATTG 2655
QY 105 rValaspheThrglnaspSerAsnTyr-----Leu***ThrGlyGlyGlnAspLysle 123
DB 2656 CTGCCATTTCACCAACAGATGATCATCTCTTACGACTGGGTCAAGGACTGCTT 2715
QY 123 uLeuArgIleTyraSpLeuAsnLysProGlnAgluProLysGluIleSerGlyHisTh 143
DB 2716 CCTCAAACTTTGGATTTGAATCAAAAGAAAGTGTGA--AAATACCATGTTTGTCAATC 2772
QY 143 rSerGlyIleLysLysAlaLeuTrpCysSerAspLysGlnIleLeuSerAlaAsp-- 162
DB 2773 AAATTCAGTCATCACTGCAATTTTCAACCATATTAAGCTTTGGCTAGTTGTTCC 2832
QY 163 -AspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeuAsnPh 182
DB 2833 TCATGGAACCTTAAAGCTTTGGATGCGACATCAACAATGAGAGCAAAAGCATTAATGT 2892
QY 182 eAsnMetSerValSerSerMetGluTyrIleProGluGluIleLeuValIleThrTy 202
DB 2893 GAAACAGTTCTCTTAATTTGGAG----- 2917
QY 202 rGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPheGluAl 222
DB 2918 -----GACCTCAAGAGCATATGGAAGT 2940
QY 222 aProAlaThrIleAsnSerAlaSer***HisProGluLysGluPheLeuValAlaGlyG 242
DB 2941 GATAGTGAAGTGTGTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 3000
QY 242 yGluAspPheLeuTyraSpLysTyraSpLysSerGlyLysGluLysGluLysSerTyry 262
DB 3001 T-----AAATCTTTTGTGGAATACAGACTCACGTTCAAAAGTGGCTGATTCGAC 3051
QY 262 sGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyLysGluLysTyry 282
DB 3052 AGGACATTTAAGTTGGTGGTGGATGATGATGATGATGATGATGATGATGATGATG 3111
QY 282 rGlySerGluAspGlyThrLeuArgLeuTrpGlnThr 294
DB 3112 ATCTTCTGATGACCAACATCAGGCTCTGGGAGACA 3148

RESULT 10
US-09-690-364-21
; Sequence 21, Application US/09690364
; Patent No. 6468795
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
; FILE REFERENCE: RTS-0190
; CURRENT APPLICATION NUMBER: US/09/690,364
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 100
; SEQ ID NO 21

```

```

: LENGTH: 7042
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (578)..(4162)
: US-09-690-364-21

Alignment Scores:
Pred. No.: 3.54e-22 Length: 7042
Score: 268.50 Matches: 77
Percent Similarity: 44.03% Conservative: 52
Best Local Similarity: 26.28% Mismatches: 128
Query Match: 14.59% Indels: 36
DB: Gaps: 7

US-09-856-836-2 (1-351) x US-09-690-364-21 (1-7042)

QY 7 Proleuthrcys-serglyhsthrargprovalasp**AlapheSerglyleth 26
Db 2365 CCGCTTAGTGTGCGCCGCCACACAGATGCTGTACCATGCGCTTCT----- 2416
QY 26 rProtyrlytyrPheleuileSeralacyslyaspGlyysPrometleuargIngl 46
Db 2417 -GAGGATGTCAGAGATAGCTTCTGTGGAGCTGATAAACCTTACAGAGTTCAAAGC 2475
QY 46 yAspThrglyaspTrrleGlythPheleuglyhstysGlyalaValatrrpGlyAlath 66
Db 2476 TGAACACAGAGAGAAATAGCTTGAAGATCAAGGCTCATGAGATGAGTGTGTGTC 2535
QY 66 rleuanslyaspAlathrlyAlaAlathrAlaAlaaspPhehrAlaLysValtr 86
Db 2536 ATTCCTACAGATGACATTTATAGCAACGCTCAGTGATTAAGGATTAAGGATTTG 2595
QY 86 pAspAlaValserGlyaspGluLeumethleu---AlahslyshsilleVallysh 105
Db 2596 GAATTCATGACTGGGGAACAGTACACACTATGATGAGCAGCAGCAAGCAATTTG 2655
QY 105 rValaspherhringlInaspSerAsnTyr-----Leu**ThrglyGlyaspLysle 123
Db 2656 CTGCCATTTCACCAACAGTACATCATCTTCTTAGCCACTGGTCAAGGACGCTT 2715
QY 123 uleuargiletyrAspLeuanslyspGluAlaGluProlysgluileSerglyhsth 143
Db 2716 CCTCAACTTGGATTTGAATCAAAAGAAATGCGA---AATACAGTCTTGGCTATAC 2772
QY 143 rserGlylelyslsAlaLeuTrrpCysSeraspLysGlnilleSerAlasp-- 162
Db 2773 AATTCAGTCAATCACTGAGATTTTCACCAAGATGATAAGCTTTGGCTAGTGTTCAGC 2832
QY 163 -AspLysThrValargLeuTrrpAspHisAlathrMethrGluValLysSerLeuansp 182
Db 2833 TGAATGAACCTTAAGCTTTGGATGCGACATCAGCAAAATGAGAGAGAAAGCTTAATGA 2892
QY 182 eAsnMetSerValSerSerMetGluTrrleProgluGlyluilleuVallethrly 202
Db 2893 GAAACAGTCTTCCTAAATTTGGAG----- 2917
QY 202 rGlyArgserlleAlaPheHisSerAlaValSerLeuGluProilleYsSerPheglua 222
Db 2918 -----GACCCCAAGAGATTTGGAAG 2940
QY 222 aProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuValaGlygl 242
Db 2941 GATAGTGAAGTGTTCGTCGTCCTGATGTCAGAGATGAATGATGATGACACAAAAA 3000
QY 242 ygluAspPheleuLyslytyrAspLysSerGlygluGluLeuGluSerlytyrly 262
Db 3001 T-----AAATCTTTTGTGGAATACACACTCAGCTTCAAGGTGGCTGATTCAG 3051
QY 262 sGlyHisPheGlyProIleHisCysValargPheSerProAspGlyluLeuutyralase 282
Db 3052 AGGACATTTAAGTGGCTTCATGCTGATGTTTCTCCTGATGAGATCATCATTTTTCAG 3111

```

```

QY 282 rGlySerGluaspGlyThrleuArgLeuTrrpGlnthr 294
Db 3112 ATCTTGTATGACCAACATCAGGCTCTGGAGACGA 3148

RESULT 11
US-09-692-508-15
: Sequence 15, Application US/09092508
: Patent No. 6291643
: GENERAL INFORMATION:
: APPLICANT: Henzel, William J.
: TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
: STREET: 3100 No. 6291643west Center, 90 South Seventh St
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/092,508
: FILING DATE: 05-JUN-1998
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 60/048,807
: FILING DATE: 05-JUN-1997
: APPLICATION NUMBER: 60/055,258
: FILING DATE: 07-AUG-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Kettelberger, Ph.D., Denise M
: REGISTRATION NUMBER: 33,924
: REFERENCE/DOCKET INFORMATION:
: TELEPHONE: 612-332-5350
: TELEFAX: 612-332-9081
: TELEX:
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7075 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: Coding Sequence
: LOCATION: 578...4192
: OTHER INFORMATION:
: US-09-092-508-15

Alignment Scores:
Pred. No.: 3.54e-22 Length: 7075
Score: 268.50 Matches: 77
Percent Similarity: 44.03% Conservative: 52
Best Local Similarity: 26.28% Mismatches: 128
Query Match: 14.59% Indels: 36
DB: Gaps: 7

US-09-856-836-2 (1-351) x US-09-092-508-15 (1-7075)

QY 7 Proleuthrcys-serglyhsthrargprovalasp**AlapheSerglyleth 26
Db 2398 CCGCTTAGTGTGCGCCGCCACACAGATGCTGTACCATGCGCTTCT----- 2449
QY 26 rProtyrlytyrPheleuileSeralacyslyaspGlyysPrometleuargIngl 46
Db 2450 -GAGGATGTCAGAGATAGCTTCTGTGGAGCTGATMAAACCTTACAGGTTCAAAGC 2508

```

```

OY 46 yAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGlyAlaValTrpGlyAlaThr 66
DB 2509 TGAACAGAGAGAGAACTCTGAAATCAAGCCTCAAGGATGAGATGCTTTGCTGTC 2568
OY 66 rLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTr 86
DB 2569 ATTCTCTACAGATGACAGATTATATAGCACTGCTCAGGGATATAAAGTAAAGATTGG 2628
OY 86 pAspAlaValSerGlyAspGluLeuMetThrLeu--AlaHisLysHisIleValLysTh 105
DB 2629 GAATTCATGACTGGGAGAACTAGTACACACTATGATGACACCTCAGCAAGCAATTCG 2688
OY 105 rValAspPheThrGlnAspSerAsnTyr-----Leu**ThrGlyGlyGlnAspLysLe 123
DB 2689 CTGCCATTTCACCAACAGTATGATCATCTCTCTTACCCACTGGCTCAAGTACGCTT 2748
OY 123 uLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluIleSerGlyHisTh 143
DB 2749 CCTCAAACTTTGGATTGAAATCAAAAGAAATGTGA---AATACCATGTTGGTCATAC 2805
OY 143 rSerGlyIleLysLysAlaLeuTyrCysSerAspLysGlnIleLeuSerAlaAsp-- 162
DB 2806 AAATTCAGTCAATCACTCAGATGATTTTCAACCAAGATGATTAAGCTTTGGCTAGTTCAGC 2865
OY 163 -AspLysThrValArgLeuTyrAspHisAlaThrMetThrGluValLysSerLeuAsnPh 182
DB 2866 TGATGAGACCTTAAAGCTTTGGATGCGACATCAACCAATATGAGAGAAACATTAATGT 2925
OY 182 eAsnMetSerValSerSerMetGluTyrIleProGluGlyIleValIleThrTy 202
DB 2926 GAAACAGTCTTCTCTAAATTTGGAG----- 2950
OY 202 rGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProIleLysSerPheGluAl 222
DB 2951 -----GACCCCTCAAGAGATATGAGAGT 2973
OY 222 aProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuValAlaGlyL 242
DB 2974 GATAGTGAAGTGTGCTGCTGCTGCGATGATGCAAGGATATGCGGACGCAAAAAA 3033
OY 242 yGluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluGluLeuGluSerTyrL 262
DB 3034 T-----AAATCTTTTGTGAAATACAGACTCAGCTCAAGTGAAGTGGCTGATTCAG 3084
OY 262 sGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyGluLeuTyrAlaSe 282
DB 3085 AGGACATTTAAGTGGGCTGCTGCTGATGATGTTTCTCTGATGATCATCATTTTGGAC 3144
OY 282 rGlySerGluAspGlyThrLeuArgLeuTyrGlnThr 294
DB 3145 ATCTTCTGATGACGACAGACATCAAGCTCTGGAGAC 3181

```

RESULT 12

```

US-09-435-115-15
; Sequence 15, Application US/09435115
; Patent No. 6346607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6346607 West Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-Aug-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ketelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.60SU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4192
; OTHER INFORMATION:
; US-09-435-115-15

```

Alignment Scores:

```

Pred. No.: 3,54e-22 Length: 7075
Score: 268.50 Matches: 77
Percent Similarity: 44.03% Conservative: 52
Best Local Similarity: 26.28% Mismatches: 128
Query Match: 14,598 Indels: 36
DB: 4 Gaps: 7

```

US-09-856-836-2 (1-351) x US-09-435-115-15 (1-7075)

```

OY 7 ProLeuThrCys-SerGlyHisThrArgProValIleAsp**AlaPheSerGlyIleThr 26
DB 2398 CCGCTTAGTGTGCTCCGCCCCACACAGATGCTGTTATCCATGCTGCTGCTTTCT----- 2449
OY 26 rProTyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLysProMetLeuArgGlnI 46
DB 2450 -GAGGATGGTCAAGAGAAATAGCTTCTTGAGGAGCTGATTAACCTTACAGCTTCAAGC 2508
OY 46 yAspThrGlyAspTrpIleGlyThrPheLeuGlyHisLysGlyAlaValTrpGlyAlaTh 66
DB 2509 TGAACAGAGAGAGAACTCTGAAATCAAGGCTCATGAGATGAGATGCTTTGTTGTC 2568
OY 66 rLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTr 86
DB 2569 ATTCTCTACAGATGACAGATTATATAGCACTGCTCAGTGGATATAAAGTAAAGATTGG 2628
OY 86 pAspAlaValSerGlyAspGluLeuMetThrLeu--AlaHisLysHisIleValLysTh 105
DB 2629 GAATTCATGACTGGGAGAACTAGTACACACTATGATGACACCTCAGCAAGCAATTCG 2688
OY 105 rValAspPheThrGlnAspSerAsnTyr-----Leu**ThrGlyGlyGlnAspLysLe 123
DB 2689 CTGCCATTTCACCAACAGTATGATCATCTCTCTTACCCACTGGCTCAAGTACGCTT 2748
OY 123 uLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluIleSerGlyHisTh 143
DB 2749 CCTCAAACTTTGGATTGAAATCAAAAGAAATGTGA---AATACCATGTTGGTCATAC 2805
OY 143 rSerGlyIleLysLysAlaLeuTyrCysSerAspLysGlnIleLeuSerAlaAsp-- 162
DB 2806 AAATTCAGTCAATCACTCAGATGATTTTCAACCAAGATGATTAAGCTTTGGCTAGTTCAGC 2865
OY 163 -AspLysThrValArgLeuTyrAspHisAlaThrMetThrGluValLysSerLeuAsnPh 182

```

```

Db 2866 TGATGACCTTAAAGCTTTGGATGCGACATGACGAAATGAGAGAAAGCATTAATGT 2925
QY 182 eanmetserValserMetgluTyrilleProgluIgluIleuValIlethrTy 202
Db 2926 GAACAGTCTTCTCTAATTTGGAG----- 2950
QY 202 rglYarSerIleAlaIhehSserAlaValSerIleuGluProIleLysSerPheGluAl 222
Db 2951 -----GACCTTCAGAGATTTGAAAGT 2973
QY 222 aProAlaThrIleasnSerAlaSer***HisProgluIugluPheLeuValAlaIglu 242
Db 2974 GATAGTGAAGTGTTCGTGCTGCTGATGCTGCAAGCATTAATGTCGACGACAAAAA 3033
QY 242 ygluAsPheIysLeuTyrIlyTyrAsPTrAsnSerGlygluIleuGluSerTyrly 262
Db 3034 T-----AAATCTTTTGTGGAATACAGACTCAGTCAAAAGGTGCTGATTCAG 3084
QY 262 sglYHisPheGlyProIleHisCysValArgPheSerProAspGlygluLeuTyrAlaSe 282
Db 3085 AGGACATTTAAGTGGTTCATGCTGATGCTGATGCTTCTTCCTGATGATCATCATTTTGAC 3144
QY 282 rglYserGluAspGlyThrIleuArgLeuTyrPgluThr 294
Db 3145 ATCTTGTGATGACGACGACATCAGGCTCTCGAGAGACA 3181

```

RESULT 13

```

US-08-188-582-17
Sequence 17, Application US/08188582
Patent No. 5534410

```

GENERAL INFORMATION:

```

APPLICANT: Tjian, Robert
APPLICANT: Comal, Lucio
APPLICANT: Dynalact, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSER: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

```

COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188, 582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435

```

ATTORNEY/AGENT INFORMATION:

```

NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

```

INFORMATION FOR SEQ ID NO: 17:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 2152 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA

```

```

FEATURE:
NAME/KEY: CDS
LOCATION: 1..2112
US-08-188-582-17

```

Alignment Scores:

```

Pred. No.: 2,666-19
Score: 238.50
Percent Similarity: 40.36%
Best Local Similarity: 25.52%
Query Match: 12.96%
DB: 1
Gaps: 11

```

```

US-09-856-836-2 (1-351) x US-08-188-582-17 (1-2152)

```

```

QY 54 ThrPheLeuGlyHisIysGlyAlaValIleTyrPgluAlaThrLeuAsnLysAspAlaThrIys 73
Db 1105 ACATTTCTCAATGCTTAAACAGAGGCTCTACTGACAGTGCATGATGATTTCTGATG 1164
QY 74 AlaAlaThrAlaAlaAspPheThrAlaLysValTyrAspAlaValSerIysAspGlu 93
Db 1165 ATTGCTGAGAGTTTGTGACAGATTCACAGTGCAGAGTGTG---TCGGTAACACCCAAAAAG 1221
QY 94 LeuMetThrLeuAlaHisIysHisIleValLysThrValAspPheThrGlnLysPserAsn 113
Db 1222 CTTCGTAGTGTCAACAGACATCAGATCTTACTTATGACAAACATCAGAT----- 1275
QY 114 TyrLeu**ThrGlyGlyIleAspLysLeuValArgIleTyrAspLeuAsnLysProglu 133
Db 1276 -----GATGCTTACAAAGATCATGAT---GAGAAAACAGCA 1311
QY 134 AlaGluProLysGluIleSerGlyHisThrSerGlyTyrIleLysAlaLeuTyrPysSer 153
Db 1312 AGTAGTGAAGATTGTTGTATGTGTACAGGCGCTTACAGACACGCTTCACTGATCG 1371
QY 154 AspAspLysGlnIleLeu---SerAlaAspAspLysThrValArgLeuTyrPAspHisAla 172
Db 1372 GATGGAACATATCTGCTTCTTCTTCCAGAGAGGAACTTGTGATTTGGAGACCTTCAA 1431
QY 173 ThrMetThr---GluValLysSerLeuAsnPheAsnMetSerValSerMetGluTyr 191
Db 1432 ACATTTACTGTTGTTGGATGATTAAGACACACACATCACTATAGGACACACATTT 1491
QY 192 IleProgluIgluIleuValIle---ThrTyrGlyArgSerIleAlaPheHisSer 210
Db 1492 TCTCCATATGGAATTAATTTGTGTACAGGGGCCATGACCGAGTAGTGGCTGGCT 1551
QY 211 AlaValSerLeuGluProIleLysSerPheGluAlaPro---AlaThrIleAsnSerAla 229
Db 1552 ACAGACCACTATACGCTTTAAGAAATTTGCCGCCATCTGCTGATTTGAAATGTAC 1611
QY 230 Ser***HisProgluLysGluPheLeuValAlaIgluIgluAspPheLysLeuTyrLys 249
Db 1612 AGATTCATCCAAATTCATTAATTTATGTTGCTACAGGGCTCTGACAGACAACTGCGGCTC 1671
QY 250 TyrAspTyrAsnSerGlygluIleuGluSerTyrTyrIysGlyHisPheGlyProIleHis 269
Db 1672 TGGACGCTCTGAATGGAATGTAAGTGTCTTACTGTGACACAGGACCAATTCAT 1731
QY 270 CysValArgPheSerProAspGlygluLeuTyrAlaSerGlySerGluAspGlyThrLeu 289
Db 1732 TCCTTGACATTTTCTCCAAATGAGATTCCTGCTACAGAGACAAACATGACAGAGTG 1791
QY 290 ArgLeuTyrGln-----ThrValValGlyLysThrTyrGly 301
Db 1792 CTTCCTTGGATTAATGACATGTTTGAATGTTGAGAAATTAAGGACACACTGATACA 1851
QY 301 ----- 301
Db 1852 GTCTGTACTTACTGATTAGTAGAGATGTAATTTTGGATCAGGTTCAATGATTAAT 1911
QY 302 -----LeuTyrLysCysVal**Pro-----GluGluAspSerGlyGlu 314

```

100

```

Patent No. 5849557
GENERAL INFORMATION:
APPLICANT: ADACHI, HIDEKI
APPLICANT: TSUJIMOTO, MASAFUMI
APPLICANT: INOUE, KEIZO
APPLICANT: ARAI, HIROYUKI
TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
TITLE OF INVENTION: AND GENE THEREOF
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER &
ADDRESSEE: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/283,917
FILING DATE: 03-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 209943/1993
FILING DATE: 03-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5849557man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2292-030-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2085 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: CDS
LOCATION: 844..2073
US-08-283-917-8
Alignment Scores:
Pred. No.: 3,17e-17 Length: 2085
Score: 221.50 Matches: 72
Percent Similarity: 44.83% Conservative: 58
Best Local Similarity: 24.83% Mismatches: 123
Query Match: 12.04% Indels: 37
DB: 2 Gaps: 9
US-09-856-836-2 (1-351) x US-08-283-917-8 (1-2085)
QY 11 SerGlyHisThrArgProValValasp***AlapheserGlyIleThrProTyrGlyTyr 30
DB 1156 AGTGTCTATGAGTCAAGTCACTGCGATTCATTTCCAT-----CCTGTGTTCACT 1206
QY 31 PhenulIleSerAlaCysAlaGlyLysProMetLeuArgGlnGlyAspThrGlyasp 50
DB 1207 GTTATGCTCTGCTCAAGAGATGCTACAAATTAAAGTGTGGATATAGAGACTGGAGAT 1266
QY 51 TrpIleGlyThrPheLeuGlyHisLysGlyAlaValTrpGlyAlaThrLeuAsnLysasp 70
DB 1267 TTGAGCAGAACTCTTAAGGGCATACAGACTGCTGTACAGATATTTCAATTCAGACCACT 1326

```

```

QY 71 AlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTrpAspAlaValSer 90
DB 1327 GGCAAGCTTCTGGCTTTCATGTTCTCAGATATGACCATTAAGCTATGGATTCAGAGGC 1386
QY 91 GlyAspGluLeuMetThrLeu---AlaHisLysHisIleValLysThrValAspPheThr 109
DB 1387 TTTGAATGATGATGAGACCATGATGCGATGACCAATGTTTCTTCTGATGACCATCATG 1446
QY 110 GlnAspSerAsnTyrLeu**ThrGlyGlnAspLysLysLeuArgIleTyrAspLeu 129
DB 1447 CCCAATGGATCATATATAGTCTGCTCCCAAGGAGGATTAACATTAATGTGGGAAGTGT 1506
QY 130 AsnLysProGlnAlaGluProLysGluLysSerGlyHisThrSerGlyIleLysAla 149
DB 1507 CAACTGCTACTGT---GTGAAGACATTCACAGACAGACAGAAATGGGTACGTATGGTGT 1563
QY 150 LeuTrpCysSerAspAspLysGlnIleLeuSer---AlaAspAspLysThrValArgLeu 168
DB 1564 CGGCCAATCAAGACGCGACTGTGATGACCACTGTTCCAAATGACCAAGACTGTGGTGT 1623
QY 169 TrpAspHisAlaThrMetThrGluValLysSerLeuAsnPheAsnMetSerValSerSer 188
DB 1624 TGGGTGTGTGCAACA---AAGCAATGCAAGGCT-----GAGCTTGAGAGA 1665
QY 189 MetGluTyrIleProGluGluIleLeuValIleThrTyrGlyArgSerIleAlaPhe 208
DB 1666 CATGACCATGTGCTAGAA----- 1683
QY 209 HisSerAlaValSerLeuGluProLysSerPheGluAlaProAlaThrIleAsnSer 228
DB 1684 -----TGCATTTCTCGGGCTCTGAAAGCTCATAT-----TCTTCATCTCTGAA 1728
QY 229 AlaSer***HisProGluLysGlu-----PheLeuValAlaGlyGly 242
DB 1729 GCAACAGATCTGAGACTAAAGTGGCAAACTGGGCCATTTCTACTGTCCGATCC 1788
QY 243 GluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluGluLeuLysTyrLys 262
DB 1789 AGGACACAGACTTCAAGATGTGGATGTCACATGCGCATGCTTMTGACCTGGTG 1848
QY 263 GlyHisPheGlyProIleHisCysValArgPheSerProAspGlyLysIleuTyrAlaSer 282
DB 1849 GGTCAATGATTAAGTGGTACGCTGAGATTCGTCTTCATTCGTGGGGGAAGTTATTTGAGT 1908
QY 283 GlySerGluAspGlyThrLeuArgLeuTyr 292
DB 1909 TGGCTGATGACAAACCTGCGGCTGTGG 1938

```

Search completed: February 6, 2003, 22:29:38
 Job time : 92 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2003, 13:11:53 ; Search time 69 Seconds

(without alignments)
2411.233 Million cell updates/sec

Title: US-09-856-836-2

Perfect score: 1840

Sequence: 1 MAMRPTLCGSHRPVVDX.....EELASNSDSIYSTPEVKA 351

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPFO_spool/US09856836/runat_31012003_144239_22239/app_query.fasta_1.519
-DB=Published.Applications_NA -OFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62
-TRANS=numa40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09856836 -ECGN_1_1_24_errunat_31012003_144239_22239
-NCPU=6 -ICPU=3 -NO_XIPYX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEM.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1347.5	73.2	1430	10	US-09-925-301-520 Sequence 520, App
2	781.5	42.5	1002	9	US-09-938-842A-510 Sequence 510, App
3	747	40.6	437	10	US-09-960-352-8644 Sequence 8644, App
4	517.5	28.1	313	10	US-09-815-343-740 Sequence 740, App

5	517.5	28.1	313	10	US-09-998-598-1947 Sequence 1947, App
6	487	26.5	311	9	US-09-920-455-140 Sequence 140, App
7	445	24.2	469	10	US-09-770-444-220 Sequence 220, App
8	425	23.1	686	10	US-09-735-705-36 Sequence 36, App
9	425	23.1	686	10	US-09-850-716A-36 Sequence 36, App
10	425	23.1	686	10	US-09-897-778-36 Sequence 36, App
11	349	19.0	361	10	US-09-998-598-2482 Sequence 2482, App
12	272.5	14.8	3025	10	US-09-764-853-29 Sequence 29, App
13	268.5	14.6	7042	10	US-09-876-667-15 Sequence 15, App
14	268.5	14.6	7072	10	US-09-876-667-15 Sequence 15, App
15	233.5	12.7	1522	9	US-10-098-841-122 Sequence 121, App
16	227.5	12.4	1378	9	US-10-098-841-121 Sequence 121, App
17	211.5	11.5	2401	10	US-09-729-674-167 Sequence 167, App
18	205.5	11.2	2151	12	US-10-042-417-1 Sequence 1, App
19	201.5	11.0	463	10	US-09-924-035A-278 Sequence 278, App
20	201.5	11.0	1410	9	US-10-109-579A-1 Sequence 1, App
21	200	10.9	262	10	US-09-923-876-3070 Sequence 3070, App
22	199.5	10.8	2551	9	US-10-098-841-285 Sequence 285, App
23	199.5	10.8	2590	9	US-10-098-841-284 Sequence 284, App
24	199.5	10.8	2710	12	US-10-044-090-646 Sequence 646, App
25	199	10.8	4285	12	US-10-044-090-646 Sequence 646, App
26	198.5	10.8	2037	10	US-09-801-368-313 Sequence 313, App
27	198	10.8	1545	10	US-09-917-800A-1690 Sequence 1690, App
28	197.5	10.7	1045	10	US-09-828-310-7 Sequence 7, App
29	196.5	10.7	1881	9	US-09-213-888-20 Sequence 20, App
30	196.5	10.7	1881	9	US-09-328-877A-20 Sequence 20, App
31	196.5	10.7	2001	9	US-09-213-888-26 Sequence 26, App
32	196.5	10.7	2001	9	US-09-328-877A-26 Sequence 26, App
33	196.5	10.7	2010	9	US-09-213-888-24 Sequence 24, App
34	196.5	10.7	2010	9	US-09-328-877A-24 Sequence 24, App
35	196.5	10.7	3550	9	US-09-328-877A-1 Sequence 1, App
36	196.5	10.7	3550	9	US-09-328-877A-1 Sequence 1, App
37	196.5	10.7	3571	9	US-09-213-888-2 Sequence 2, App
38	196.5	10.7	3571	9	US-09-328-877A-2 Sequence 2, App
39	195.5	10.6	1211	9	US-10-119-932-2 Sequence 2, App
40	194	10.5	1288	10	US-09-770-445-18 Sequence 18, App
41	192	10.4	2369	10	US-09-908-805B-20 Sequence 20, App
42	191	10.4	1089	10	US-09-917-800A-1444 Sequence 1444, App
43	191	10.4	1093	10	US-09-969-708-214 Sequence 214, App
44	191	10.4	1093	10	US-09-880-107-2302 Sequence 2302, App
45	187	10.2	2186	9	US-09-943-689A-1 Sequence 1, App

ALIGNMENTS

RESULT 1
US-09-925-301-520
Sequence 520, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antipodules
FILE REFERENCE: PA106
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: US/09/925,301
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 520
LENGTH: 1430
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (104)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature
LOCATION: (105)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature

LOCATION: (1428)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-520

Alignment Scores:

Pred. No.:	4,28e-154	Length:	1430
Score:	1347.50	Matches:	257
Percent Similarity:	96.31%	Conservative:	4
Best Local Similarity:	94.83%	Mismatches:	9
Query Match:	73.23%	Indels:	1
DB:	10	Gaps:	1

US-09-856-836-2 (1-351) x US-09-925-301-520 (1-1430)

```
OY 81 petherAlaAlaValTTPspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 39 TTCACAGCCAAAGTGGGATGCTGCTCAGAGATGAAATGATGACCTGCTCATTA 98
OY 101 HisLeuValLysThrValAspPheThrGlnAspSerAsnTyrLeu**ThrGlyGlyGln 120
DB 99 CACATNNCTCAAGACTGGATTCACGAGAGATGAAATTAATTTGTTAAACCGGGGACAG 158
OY 121 AspLysLeuLeuArgLleTyrAspLeuAsnLysProGluAlaGluProLysGluLleSer 140
DB 159 GATTAACCTTTACGATATATGACTTGAACAAACCTGAAGCAGAACTTAAGAAATTAGT 218
OY 141 GlyHisThrSerGlyLleLysLysAlaLeuTTPCysSerAspAspLysGlnLleLeuSer 160
DB 219 GGTCACTACTCTGCTGATTAATAAAAGCTCTGCTGACGAGATTAACAGATTCCTTCT 278
OY 161 AlaAspAspLysThrValArgLeuTTPAspHisAlaThrMetThrGluValLysSerLeu 180
DB 279 GCTGATGACAAAACCTTTGGACTTTGGATCAGCTACATGACAGAGAAATCTCTA 338
OY 181 AsnPheAsnMetSerValSerSerMetGluTyrLleProGluGluLleLeuValLle 200
DB 339 AATTAAATATGCTCTGTAGTATGATGAAATATTCCTGAGGAGAGATTTGGTTATA 398
OY 201 ThrTyrGlyArgSerLleAlaPheHisSerAlaValSerLeuGluProLleLysSerPhe 220
DB 399 ACTTAAGAGAGATCTATGCTTTTCATGCTGAGTAAAGTTTGGACCAATTAATCTTTT 458
OY 221 GluAlaProAlaThrLleAsnSerAlaSer**HisProGluLysGluPheLeuValAla 240
DB 459 GAAGCTCCGCAACCATTCATCTGATCTTCATCCCTGAGAAAGAAATTTCTTGCA 518
OY 241 GlyGlyGluAspPheLysLeuTyrLysTyrAspTyrAsnSerGlyGluGluLysSer 260
DB 519 GGGGGTGAAGATTTTAACTTTAAGTATGATTAATATGATGAGAGAAATTAAGATTC 578
OY 261 TyrLysGlyHisPheGlyProLleHisCysValArgPheSerProAspGlyLleLeuTyr 280
DB 579 TACAAGGACACTTTGCTCTATCTACTGCTGATGATTTAGTCTGATGGAACCTCTAT 638
OY 281 AlaSerGlySerGluAspGlyThrLeuArgLeuTTPGlnThrValValGlyLysThrTyr 300
DB 639 GCCAGTGGTCAAGAAAGATGAGACATGAGCAAACTGCTGTAAGGAAACCTAT 698
OY 301 GlyLeuTyrLysCysVal**ProGluGluAspSerGlyGluLeuAlaLysProLysLle 320
DB 699 GGGCTTTGGAATGTGCTCTGGAAGAAATGATGATGCTGCTGGAAGAAACCTAAGTTT 758
OY 321 GlyPheProGluThrAlaGluGluGluLeuAlaGluGluLleAlaSerGluAsnSerAsp 340
DB 759 GGTTCCTCAAGACAAAGAGAGAGAGCTA---GAAGAATTTGCTTCAAGAAATTCAGAT 815
OY 341 SerLleTyrSerSerThrProGluValLysAla 351
DB 816 TGCATCTTCTCTCAAGCTCTGATGATTAAGGCC 848
```

RESULT 2

US-09-938-842A-510
Sequence 510, Application US/09938842A

```
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIP300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 510
LENGTH: 1002
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-510
```

Alignment Scores:

Pred. No.:	1.46e-85	Length:	1002
Score:	781.50	Matches:	148
Percent Similarity:	66.02%	Conservative:	56
Best Local Similarity:	47.90%	Mismatches:	94
Query Match:	42.47%	Indels:	11
DB:	9	Gaps:	4

US-09-856-836-2 (1-351) x US-09-938-842A-510 (1-1002)

```
OY 7 ProLeuThrCysSerGlyHisThrArgProValValAsp**AlaPheSerGlyLleThr 26
DB 10 CCGTATGAGTGGCCAGGCGCAATTCCTGCTCCCGCTGATGCTGCTTATCCGGTACT 69
OY 27 ProTyrGlyTyrPheLeuLleSerAlaCysLysAspLysProMetLeuArgGlnGly 46
DB 70 CCAGATGGTCTCTTTCATTAAGCCAGCAAGATTCGAATGAGATGAGAAATGA 129
OY 47 AspThrLysAspTTPLleGlyThrPheLeuGlyHisGlyAlaValTTPGlyAlaThr 66
DB 130 GAGCTGTGACGATGGATGGACTTTGAAGACATAGGAGAGAGATTTGGAGTTGAGC 189
OY 67 LeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTTP 86
DB 190 CTGATTAATAATGCTATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
OY 87 AspAlaValSerGlyAspGluLeuMetThrLeuAlaHisLysHisLleValLysThrVal 106
DB 250 AATGCAATTCAGAGAGATGAAATGCACTCTTGAACCAAGCAATTCCTGCTGCTGCT 309
OY 107 AspPheThrGlnAspSerAsnTyrLeu**ThrGlyGlyGlnAspLysLeuLeuArgLle 126
DB 310 GCCTTTCTGAGACCTACCCGTTTACTCACTGCTGCAATGGAATAATTAATCTTCGATA 369
OY 127 TyrAspLeuAsnLysProGluAlaGluProLysGluLleSerGlyHisThrSerGlyLle 146
DB 370 TTCATTTGAATGCGCCAGAGACGACCTCCAAAGAGCGGAATTCCTGCTGCTGCTGCT 429
OY 147 LysLysAlaLeuTTPCysSerAspAspLysGlnLleLeuSerAla-----AspAspLys 164
DB 430 AGAAGCTGCAATGCTTCATAGTATGATTAATCAATTCCTTAAAGCTCTTGCACAGATCCGGT 489
OY 165 ThrValArgLeuTTPAspHisAlaThrMetThrGluValLysSerLeuAsnPheAsnMet 184
DB 490 GACATTTAGTATGAGATTAAGAGAGAGAGATGCTTCAATTAATTAAGAAACAAGTCC 549
OY 185 SerValSerSerMetGluTyrLleProGluGluLleLeuValLleThrTyrGlyArg 204
DB 550 CCAATTACTAGTGAAGTAAGTCAAGATGGGAGATCACTACTGCTGATGATGATCT 609
```


Alignment Scores:	6.4e-82	Length:	437
Pred. No.:	747.00	Matches:	141
Score:	99.31%	Conservative:	3
Percent Similarity:	97.24%	Mismatches:	1
Best Local Similarity:	40.60%	Indels:	0
Query Match:	10	Gaps:	0
DB:			
US-09-856-836-2 (1-351) x US-09-960-352-8644 (1-437)			
QY	138	GIUIESerGIhSIrSIerGIrIeLSySaIaLeuTrPyCSerSaspDysGIh	157
Db	3	GAATACATGGCTCACCCCTCGTATTAAAAAGCCTTGTTGGTGCATGAGGATAAACAG	62
QY	158	IIleuSerSaIaaspDysTrValrIgreuTrpSpHsIaTrmLeTrGluVal	177
Db	63	ATCCTTTAGCGGATGATATAAAGCTGTCGCCCTTTGGGATCATGCTACTATAGCGAAGTG	122
QY	178	LYSSerLeuAnpHeaSmelSeVasIerSeRmetGluTrIleProGIuGIyGluIle	197
Db	123	AAATCTCAATTTTAAATATATCCGTAGCGATGATGCATCTCCGAGGAGAGAAATT	182
QY	198	LeuValIIleTrIryGIyargSerIIleIaphelISerSaIaVasIerLeuGIuProIle	217
Db	183	TTGGTAATACCTTATAGCGATCATATCTTTCAATATGAGAGTAAGTTGGACCAATT	242

Alignment Scores:		
Pred. No.:	2,945-54	Length: 313
Score:	517.50	Matches: 97
Percent Similarity:	94.29%	Conservative: 2
Best Local Similarity:	92.38%	Mismatches: 5
Query Match:	28.12%	Indels: 1
DB:	10	Gaps: 1

QY	246	LysLeuTYrLysTYrAspTYrAspSerGlyGluGluLeuGluSerTYrLysGluHisPhe	265
Db	1	AAACCTTAAAGTAGATATTAAATAGTGGAGAAGAAATTAAGAACTCTCAAGAGGACACTTT	60
QY	266	GlyProLeuHisCysValArgPheSerProAspGlyGluLeuTYrHisAspGlySerGlu	285
Db	61	GCTCTATTCACCTGGTAGATTAGTCTCATAGAGAAACTCTATGCCAGAGGTTCAGAA	120
QY	286	AspGlyThrLeuArgLeuTYrGlnThrValValGlyThrTYrGlyLeuTrpLysCys	305
Db	121	GATGGAACATTGCAGCTATGGCAAACTGTGGAGAAAACGTAATGGCTTTGGAAATGT	180
QY	306	Val***ProGluGluAspSerGlyGluLeuAlaLysProLysIleGlyPheProGluThr	322
Db	181	GGGCTTCCTGAAGAAGATAGTGAGTGGCAAAAGCCAAAGATGGTTCTCCAGAGCA	240
QY	326	AlaGluGluGluLeuAlaGluGluLeuLeuAspGluLysSerAspSerIleTYrSerSer	345
Db	241	ACAGAGAGAGGACCTA---GAAGAAATTCCTTCAGAGGAAATTCAGATTCATCTTCCTCA	299
QY	346	ThrProGluLysLys 350	
Db	298	GCTCCTGAGTGAAG 312	

RESULT 5
US-09-998-598-1947

Sequence 1947, Application US/0998598
Patent No. US20020150922A1
GENERAL INFORMATION:
APPLICANT: Stolk, John A.
APPLICANT: Xu, Jiangchun
APPLICANT: Chenault, Ruth A.
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE OF INVENTION: DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.561
CURRENT APPLICATION NUMBER: US/09/998,598
CURRENT FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 2606
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 1947
LENGTH: 313
TYPE: DNA
ORGANISM: Homo sapiens
US-09-998-598-1947

Alignment Scores:
Pred. No.: 2,948-54 Length: 313
Score: 517.50 Matches: 97
Percent Similarity: 94.29% Conservative: 2
Best Local Similarity: 92.38% Mismatches: 5
Query Match: 28.12% Indels: 1
Gaps: 10

US-09-856-836-2 (1-351) x US-09-998-598-1947 (1-313)

QY 246 LysLeuTyrLysTyrAspTyrAsnSerGlyGluGluSerTyrLysGlyHisPhe 265
DB 1 AACTTTAATAGTATATATATAGTGGAGAAATAGATCTTCAAGGACACTTT 60
QY 266 GlyProIleHisCysValAlaPheSerProAspGlyGluLeuTyrAlaSerGlySerGlu 285
DB 61 GGTCTATTCATCTGTGTGATTTAGTCTCTGATGAGAACTTATGCCAGGCTTCA 120
QY 286 AspGlyThrLeuArgLeuTyrPglThrValAlaGlyThrTyrGlyLeuTyrPlyCys 305
DB 121 GATGGACATTTGAGACTATGGCAACTGTGTAGGAAAACGTATGGCTTGGAAATGT 180
QY 306 Val***ProGluGluAspSerGlyGluLeuAlaLysProLysIleGlyPheProGluThr 325
DB 181 GTGCTTCTGAAAGAGATGTGTGAGCTGGCAAGCCAAAGATTGTTTCCAGAGCA 240
QY 336 AlaGluGluGluLeuAlaGluGluLeuAlaSerGluAsnSerAspSerIleTyrSerSer 345
DB 241 ACAGAAAGAGAGCTA---GAAGAAATTCCTTCAGAGATTCAGATTGCATCTTCCCTCA 297
QY 346 ThrProGluValLys 350
DB 298 GCTCCTGATGTTAAG 312

RESULT 6
US-09-920-455-140
Sequence 140, Application US/09920455
Patent No. US20020168647A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE OF INVENTION: AND DIAGNOSIS OF HEAD AND NECK CANCER
FILE REFERENCE: 210121.540
CURRENT APPLICATION NUMBER: US/09/920,455
CURRENT FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 275
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 140
LENGTH: 311
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: misc-feature
LOCATION: 115, 199, 201, 235, 238
OTHER INFORMATION: n = A,T,C or G
US-09-920-455-140

Alignment Scores:
Pred. No.: 1,488-50 Length: 311
Score: 487.00 Matches: 90
Percent Similarity: 95.79% Conservative: 1
Best Local Similarity: 94.74% Mismatches: 4
Query Match: 26.47% Indels: 0
Gaps: 0

US-09-856-836-2 (1-351) x US-09-920-455-140 (1-311)

QY 212 ValSerLeuGluProIleLysSerPheGluAlaProAlaThrIleAsnSerAlaSer*** 231
DB 23 GTAAATTGGACCCCAATTAATCTTTGAAGCTCTGCAACCAATCAATTCGATCTCTT 82
QY 232 HisProGluLysGluPheLeuValAlaGlyGlyLysAspPheLysLeuTyrLysTyrAsp 251
DB 83 CATCTGAGAAAGAAATTTCTTGTGACGGCGNAGATTTTAACTTTATAGTATGAT 142
QY 252 TyrAsnSerGlyGluGluLeuGluSerTyrLysGlyHisPheGlyProIleHisCysVal 271
DB 143 TATATAGTGGAGCAAGAAATTAAGATCTTCAAGACACATTTGGGCTTATTCATGNGC 202
QY 272 ArgPheSerProAspGlyGluLeuTyrAlaSerGlySerLysAspGlyThrLeuArgLeu 291
DB 203 AGATTAGTCTGTAGGAGAACTTATGCGACGNGNCAAAATAGGAACTTGAAGACTA 262
QY 292 TrpIleThrValAlaGlyLysThrTyrGlyLeuTyrPlyCysVal 306
DB 263 TGGCAAACTGTGTAGGAAAACGTATGGCTTGGAAATGTGTG 307

RESULT 7
US-09-770-444-220

Sequence 220, Application US/09770444

Patent No. US2002023280A1

GENERAL INFORMATION:

APPLICANT: Gorlach, Jorn
APPLICANT: An, Yong-Olang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Moessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE OF INVENTION: Thailanda
FILE REFERENCE: 2027 (PARA-016PRV)
CURRENT APPLICATION NUMBER: US/09/770,444
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,502
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 220
LENGTH: 469
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-444-220

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3.52e-45	445.00	74.15%	55.78%	24.18%	469	82	27	38	0	0

US-09-856-836-2 (1-351) x US-09-770-444-220 (1-469)

```

OY 15 ACPGVALAASP**AlapheserGlyIethrProTyGlyTyrePhleuIeser 34
DB 13 AACCTGTGTGATTTGTTCTACACGACATCATCTCTGATGTTCTCTCTCATACAC 72
OY 35 AAcysLysAspGlyLysPrometleuArGInGlyAspPthGlyAspPthIleGlyThr 54
DB 73 GCAAGTAAAGATTCTCAACCAATGTTGAAATGGGAAATGGAGATGATGATGTGACA 132
OY 55 PheleuGlyHisLysGlyAlaValITrpglyAlaThrLeuAsnLysAspAlaThrLysAla 74
DB 133 TTTGAAGGTCAATAAGGTCTGTGTGAGAGTCTGTCTGTGATTAACAATGCTTACGTCA 192
OY 75 AAlaThrAlaAlaAlaAspPheThrAlaLysValITrpaAlaValISerGlyAspGluLeu 94
DB 193 GCTTGTGATCTGCTGATTTTTCACGGAAGCTTTGGATGCTTACTGAGGATGTCTTG 252
OY 95 MetThrLeuAlaHisLysHisLysIleValLysThrValAspPheThrGlnAspSerAsnTy 114
DB 253 CATCTTTTGGACCAACAGATATTTGTGAGCATGCGGCTTCTCAGAGATACGAATAT 312
OY 115 Leu***ThGlyGlyLysAspLysLeuLeuArgIleTyAspLysAsnLysProGluAla 134
DB 313 CTAATCACAGGAGGATTGAGAAATTTCTTCGTTTTCAGCTTAATTCGCTTGCATGCA 372
OY 135 GluProLysGluIleSerGlyHisThrSerGlyLysLysAlaLeuITrpsSerasp 154
DB 373 CCTCTACAGAAATGATTAATCTCTGTTCTATACAGAACCTAACATGCTTACGCT 432
OY 155 AspLysGlnIleLeuSerAla 161
DB 433 GATCAACAAATATTAAGTCT 453

```

RESULT 8

US-09-735-705-36

Sequence 36, Application US/09735705

Patent No. US20020052329A1

GENERAL INFORMATION:

APPLICANT: Wang, Tonglong

APPLICANT: Fan, Liqun

APPLICANT: Kalos, Michael D.

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Hosken, Nancy

APPLICANT: Fanger, Gary R.

APPLICANT: Li, Samuel X.

APPLICANT: Wang, Aljun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Henderson, Robert A.

APPLICANT: McNeill, Patricia D.

APPLICANT: Fanger, Neil

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C14

CURRENT APPLICATION NUMBER: US/09/735.705

CURRENT FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 419

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 36

LENGTH: 686

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(686)

OTHER INFORMATION: n = A,T,C or G

US-09-735-705-36

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1.69e-42	425.00	83.97%	81.68%	23.10%	686	107	3	21	7	0

US-09-856-836-2 (1-351) x US-09-735-705-36 (1-686)

```

OY 1 MetAlaMetArgGlnIThrProLeuIThrCysSerGlyHisIThrArgProValAASP*** 20
DB 277 ATGGCAATNAGACAGAGCCGCTCACCCTGCTGGCAGACAGCCGCTGATTTG 336
OY 21 AlapheserGlyIethrProTyGlyTyrePhleuIleSerAlaCysLysAspGlyLys 40
DB 337 GCTTCACTGCGCATCAC-CCATTATGGGATTTCTTAATCAGCCGCTGCAGAGATGTTAA 395
OY 41 PrometleuArGInGlyAspPthGlyAspPthIleGlyThrLeuGlyHisLysGly 60
DB 396 CCAATGCTACGCCAGGAGATTAACAGAGACTGATGAGACATTTTGGGCTCTAAAGGT 455
OY 61 AAlaValITrpglyAlaThrLeuAsnLysAspAlaThrLysAlaAlaIThrAlaAlaAsp 80
DB 456 CT-GTTTGGGTGCAACACTAATTAAGATGCCACCAAGCAGCTACAGCCTGCAGAT 514
OY 81 PheThrAlaLysValITrpaAlaValISerGlyAspGluLeuMetThrLeuAlaHisLys 100
DB 515 TTCACAGCCCAAGTGTGGATGCTGCTCAGAGANATTAATGATTAACCTG-GCTCATTA- 572
OY 101 HisIleValLysThrValAspPheThrGlnAspSerAsnTyLeu***ThGlyGlyGln 120
DB 573 CACATGTCAAGATGAGGATTTCC-CAGATAT-TATTATTTGTTTACCGGGGAGNAG 630
OY 121 AspLysLeuLeuArgIleTyAspLysAla 131
DB 631 GATTA-CTGTTTCCNCTATTTTAATTCACAAA 662

```

RESULT 9

US-09-850-716A-36

Sequence 36, Application US/09850716A

Patent No. US20020115139A1

GENERAL INFORMATION:

APPLICANT: Kalos, Michael D.

APPLICANT: McNeill, Patricia D.

APPLICANT: Retter, Marc W.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C15

CURRENT APPLICATION NUMBER: US/09/850.716A

CURRENT FILING DATE: 2001-05-07

NUMBER OF SEQ ID NOS: 440

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 36

LENGTH: 686

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(686)

OTHER INFORMATION: n = A,T,C or G

US-09-850-716A-36

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:
1.69e-42	425.00	83.97%	81.68%	23.10%	686	107	3	21	7


```

?      TYPE: nucleic acid
?      STRANDEDNESS: double
?      TOPOLOGY: linear
?
?      MOLECULE TYPE: cdna
?
?      FEATURE:
?      NAME/KEY: Coding Sequence
?      LOCATION: 578...4159
?      OTHER INFORMATION:
?      SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-876-667-1

```

Assignment: Decades:	
Pred. NO.:	6.15e-22
Score:	266.50
Percent Similarity:	44.03%
Best local Similarity:	26.28%
Query Match:	14.59%
DB:	10
Length:	7042
Matches:	77
Conservative:	52
Mismatches:	128
Indels:	36
Gaps:	7

US-09-856-836-2 (1-351) x US-09-876-667-1 (1-7042)

Oy	7	Prneuhncrcys-SerclylshsthrnprvovAlaValasp***AlaphSerSgylleth	26
Db	2365	CCGGTtAGTtGTCCGCCCCCACCACAGATGCTGTtTACCACTGCCTCTTTCt-----	2418
Oy	26	rProtyrGlytYrPheleuileSerAlaacylsyaspGlyLysPrometleuArgIncl	46
Db	2417	-GAGGATGCTCAGAGATTAACCTCTTGAGAGTAAACCTTACAGTGTTCAAAGC	2475
Oy	46	yAsPthGlyAspPrIleclYthrPheleuGlyHslySgylAlaValtPrpAlaLth	66
Db	2476	TGAACACGAGAGAAACCTTGTACAAATCAAGGCTCAAGAGATGAAGTGTtTGTGTGC	2533
Oy	66	rLeuAspLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValtyr	86
Db	2526	ATTCTCTACAGATACACAGATTTATAGCAACCTCTCAGTGGATAAACCAAGTGAACATTTG	2599
Oy	86	pAspAlaValSerGlyAspGluLeuMetThrLeu--AlaHsLysHsIleValLysTh	105
Db	2596	GAATTCATGACTGCTGGGAGACTAGACACACCTTATGATGAGCACTCAGACAGTCAATTTG	2655
Oy	105	rValAspPheThrLnsPheLsrAsnTy-----Leu***ThrGlyGlyLnsAspLysIe	123
Db	2656	CTGCCATTTCACCAACAGTGTGCTATCTCTTCTTACGCCACTGGGTCAAGTGAAGCTGT	2715
Oy	123	uLeuArgIleTyrrPheLysLysProGluAlaGluPrGlySgylIleSerGlyHsIth	143
Db	2716	CCTCAACCTTTGGATTTGAATCAAAAAGAAATGTCCA---AATACATGTTTGGTCAATAC	2774
Oy	143	rSerGlyIleLysLysAlaLeuTrpCysSerAspAspLysGlnIleLeuSerAlaasp--	162
Db	2773	AAATTCAGTCAATCAGTCAGATTTTCCACAGATGATTAAGCTTTTGGTATATGTTTACAC	2833
Oy	163	AspLysThrValArgLeuTrpAspHsIalArMetThrGluValLysSerLeuAsnPh	182
Db	2833	TGATGGAACCTTAAAGCTTTGGGTGGCATGACAAATGAAGAGAAAGCAATTAATCT	2893
Oy	182	eAsnMetSerValSerSerMetGluTyrlIeProGluGlyuIleLeuValIleThTy	202
Db	2893	GAACACGTTCTTCTAAATTTGGAG-----	2911
Oy	202	rGlyArgSerIleAlaPheHsIserAlaValSerLeuGluProIleLysSerPheGluAl	222
Db	2918	-----GACCTCAAGAGATATAGAACT	2944
Oy	222	aProAlaThrIleAsnSerAlaSer***HisProGluGlySglnPheLeuAlaLaglycl	242
Db	2941	GATAGTGAAGTGTtTCTTCGCTGCTGCTGATGTCGCAAGAGATATAGTGGACGCAAAAAA	3000
Oy	242	yGluAspPheLysLeuTyrlLysTyrrAsnSerGlyGluGluLeuGluUserTytyl	262
Db	3001	T-----AAATTCTTTTTGTGGAATACAGACTCAGTTCAAAGGTGGCTGATTTGCAG	3055

Oy	262	scgIlysPhgclgylprollethscysvalrtrpPheSerProAspGlyGlnLeuTyraIase	282
		: : : :	
Db	3052	AGGACATTTAAAGTGGGTTCATGGTGGATGATTTTCTCCATGATGATCATATTTTGAC	3111
Oy	282	rcGlySerGluAspGlyThrLeuArgLeuArgpGlnTrp	294
		: : : : :	
Db	3112	ATCTTTCGATGACACAGCAATTCAGGCTCTGGAGAGACA	3148

RESULT 14
HE-00-076

US-09-876-667-15
; Sequence 15, Application US/09876667
Patent No. US20020107370A1

SERIAL INFORMATION:
 APPLICANT: Henzel, William J.
 TITLE OF INVENTION: APA-1, AN ACTIVATOR OF C ASPASE-3
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 Dr. William J. Henzel
 Department of Chemistry
 University of California
 1113 Chemistry Building
 University of California
 San Diego
 La Jolla, CA 92037

ADDRESSEE: Merchante, Gould, Smith, Edell, Weiler & Schmidt
STREET: 3100 No. US20020107370A West Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA: -----

APPLICATION NUMBER: US/09/876,667
FILING DATE: 07-Jun-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/435,115
FILING DATE: <Unknown>
APPLICATION NUMBER: 60/055,258
FILING DATE: 07-AUG-1997

ATTORNEY/AGENT INFORMATION:
NAME: Kettelberger, Ph.D., Denise M
REGISTRATION NUMBER: 33,924
REFERENCE/DOCKET NUMBER: 11669.6US01
TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
TELEX: <Unknown>

RMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 7075 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: Coding Sequence

LOCATION: 578...4192
OTHER INFORMATION:
SEQUENCE DESCRIPTION: SEQ ID NO: 15

Case No.	Case Name	Case Type	Case Status	Case Date
6-667-15	6-667-15	6-667-15	6-667-15	6-667-15

Alignment Scores:	
Pred. No.:	6.19e-22
Score:	268.50
Percent Similarity:	44.03%
Best Local Similarity:	26.28%
Query Match:	14.59%
DB:	10
	7
Length:	7075
Matches:	57
Conservative:	52
Mismatches:	128
Indels:	36
Gaps:	7

US-09-856-836-2 (1-351) x US-09-876-667-15 (1-7075)

QY 7 ProLeuThrcys-SerGlyHisThraGrProValValasp***AlaPheSerGlyLeu 26
|||||:||||| ||||| ||||| |||||

Db 2398 CGCGTACTTGTGCGCCGCCACACAGATGCTGTTTACATGCTGTTTTCT----- 2443

```

QY 26 rProtyrGlyTyrPheLeuIleSerAlaCysLysAspGlyLysPrometLeuArgGlnG 46
Db 2450 -GAGATGGTCAAGAAATGCTTCTGTGGAGCTGATAAACCTTACAGAGTGTCAAGC 2508
QY 46 yAspThrGlyAspThrIleGlyTyrPheLeuGlyHisLysGlyAlaValTrpGlyAlaThr 66
Db 2509 TCAACAGAGAGAAACTTCTGAAATCAAGAGCTATGAGTGAAGTCTTGTGTTC 2568
QY 66 rLeuAsnLysAspAlaThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTrp 86
Db 2569 ATCTCTACAGATGACAGATTTATAGCAACCTGCTCAGTGATTAATAAGTAAATTTG 2628
QY 86 pAspAlaValSerGlyAspGluLeuMetThrLeu---AlaHisLysHisLysValLysThr 105
Db 2629 GAATCTATGACTGGGAGGATGATGACACACCTATGATGAGCCTCAGCAAGCAATGTTG 2688
QY 105 rValAspPheThrGlnAspSerAsnTyr-----Leu**ThrGlyGlyGlnAspLysLe 123
Db 2689 CTGCCATTTCAACACAGTACATGATCTTCTTACCCACTGGGTCAAGTCACTGCTT 2748
QY 123 uLeuArgIleTyrAspLeuAsnLysProGluAlaGluProLysGluLysSerGlyHisThr 143
Db 2749 CTTCAAACTTTGGGATTTGAAATCAAAAAGAAATGTCGA---AATACCATGTTGGTCAATC 2805
QY 143 rSerGlyIleLysLysAlaLeuTyrCysSerAspAspLysGlnIleLeuSerAlaAsp-- 162
Db 2806 AAATTCATCATATCATCTCAATTTTCCACAGATGATTAAGCTTTGGCTTGTTCAGC 2865
QY 163 -AspLysThrValArgLeuTrpAspHisAlaThrMetThrGluValLysSerLeuAsnPh 182
Db 2866 TGAATGGAACCTTAAAGCTTGGATGCCACATCAGCAAAATGAGAGCAAAACCATTAATGT 2925
QY 182 eAsnMetSerValSerSerMetGluTyrIleProGluGlyLysIleLeuValIleThrTy 202
Db 2926 GAACACGTTCTTCTTAAATTTGGAG----- 2950
QY 202 rGlyArgSerIleAlaPheHisSerAlaValSerLeuGluProLysSerPheGluAl 222
Db 2951 -----GACCTCAAGAGATATGAGAGT 2973
QY 222 aProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuValAlaGlyL 242
Db 2974 GATAGTCAAGCTGTTGCTGCTGCTGCTGATGTCGAAGATGATGATGATGATGATGATG 3033
QY 242 yGluAspPheLysLeuTyrTyrAspTyrAsnSerGlyGluGluLeuLeuSerTyrLy 262
Db 3034 T-----AAATCTTTTGGATATGACACTCAGCTTCAAGAGTGGCTGATTCAG 3084
QY 262 sGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyLysLeuTyrHisL 282
Db 3085 AGGACATTTAAGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3144
QY 282 rGlySerGluAspGlyTyrPheLeuArgLeuTrpGlnThr 294
Db 3145 ATCTCTGATGACACAGCAATCAAGGCTGTGGAGACA 3181

```

```

RESULT 15
US-10-098-841-122
Sequence 122, Application US/10098841
Publication No. US20020197679A1
GENERAL INFORMATION:

```

```

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Duntui
APPLICANT: Wang, Zhilwei

```

```

APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
FILE REFERENCE: 784CJP2
CURRENT APPLICATION NUMBER: US/10/098,841
PRIOR APPLICATION NUMBER: 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
NUMBER OF SEQ ID NOS: 331
SOFTWARE: PL-FL-Genes Version 1.0
SEQ ID NO 122
LENGTH: 1522
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (456)..(1061)
US-10-098-841-122

```

Alignment Scores:

```

Pred. No.: 1.06e-18 Length: 1522
Score: 233.50 Matches: 77
Percent Similarity: 40.18 Conservative: 54
Best Local Similarity: 23.628 Mismatches: 130
Query Match: 12.69% Indels: 65
DB: Gaps: 12

```

US-09-856-836-2 (1-351) x US-10-098-841-122 (1-1522)

```

QY 52 rLeGlyThrPheLeuGlyHisLysGlyAlaValTrpGlyAlaThrLeuAsnLysAspAla 71
Db 39 CTGGTACGAGCTCGAGTCCAGTACTAGTGTGGATTCAG----- 86
QY 72 ThrLysAlaAlaThrAlaAlaAspPheThrAlaLysValTrp---AspAlaValSer 90
Db 87 -----ACTGCCGATCTGGGTACCCAAATGCAAA 116
QY 91 gLysAspGluLeuMetThrLeuAlaHisLysHisLysValLysThrValAspPheThrGln 110
Db 117 GGTGAGTCCAGCTGTGTTGTGACACACAGCAGATGAGAGTCCACTTCTGCACT 176
QY 111 AspSerAsnTyrLeu**ThrGlyGlyGlnAspLysLeuLeuArgIleTyrAspLeuAsn 130
Db 177 GATGGCAGTCTCTGCTGACAGCTCTGACGACAGACAGTCAAGTGTGGCACTCAT 236
QY 131 LysProGluAlaGluProLysGluLysSerGlyHisThrSerGlyIleLysLysAlaLeu 150
Db 237 CCC---CAAAATTCCTGTTCTCCCTGAGCCAGCATATCACTGGCTGCCCAAG 293
QY 151 TrpCysSerAspAspLysGlnIleLeuSerAla---AspAspLysThrValArgLeuTrp 169
Db 294 TTTCTCCCCGAGCGGCGGCTCATGCTGTGCAATGATGACAGACTGTTAAGCTGTGG 353
QY 170 AspHisAlaThrMetThrGluValLysSer---LeuAsnPheAsnMetSerValSerSer 188
Db 354 GACAAAGAGAGCGGGGAATGTGTCCACCTGATTTGTGACATGGCGGCTTGTCACTAT 413
QY 189 MetGluTyrIleProGluGlyLysLeuValIleThrTyrGlyArgSerIleAlaPhe 208
Db 414 GTGACATTCACCCAGCGGAGGAGTGCATGCC-----GCTGCCGAGATG 458
QY 209 HisSerAlaValSerLeuGluProIleLysSerPheGluAla----- 222
Db 459 GACACACAGTGAAGTGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTG 518
QY 223 --ProAlaThrIleAsnSerAlaSer**HisProGluLysGluPheLeuValAlaGly 241

```

```

Db      519  CACAGTGCAGCAGTGAACGGCGCTCTCTTCCACCCCTGGGAAACTACTGATCACAGCC 578
           |||  ::|||  |||  |||||  ::|||::
QY      242  GYGLIAspPheLeuTyrLysTyrAspTyrAsnSerGlyGluGluLeuGluSerTyr 261
           |||  |||  |||  |||  |||  ::
Db      579  TCCAGTGACTCAACCTGAAGATCCTGGACCTGATGAGAGGCGGCTGCTACACACTC 638
           |||  |||  |||  |||  |||  ::
QY      262  LysGlyHisPheGlyProIleHisCysValArgPheSerProAspGlyGluLeuTyrAla 281
           |||||  |||||  |||  |||||  |||||  ::|||
Db      639  CACGGGATCAGGGAGCACCCACTGTGCTTTCAGAAACGGGGAGATATTTTGT 698
           |||||  |||||  |||  |||||  |||||  ::|||
QY      282  SerGlySerGluAspGlyThrLeuArgLeuTrpGlnThr-----ValValGlyLys 298
           |||||  |||  |||  ::|||  ::|||
Db      699  TCTGAGGCTCTGATGAACAAGTATGTTGGAAGAGTACTTGATATGTT----- 752
           |||||  |||  |||  ::|||  ::|||
QY      299  ThrTyrGlyLeuTrpLysCysVal***ProGluGluAspSerGlyGluLeuAlaLys--- 317
           |||  |||  |||||  ::|||
Db      753  -----GATCATGAGAAAGTCAAGAAAGTG 776
           |||  |||  |||||  ::|||
QY      318  -----ProLysIleGlyPhePro 323
           |||  |||  |||||  ::|||
Db      777  CCGAGGCCCCAGCCACACTGGCCAGCTCCATGGGGAATCTGCCAGAGATGGACTTCCT 836
           |||  |||  |||  |||  |||  ::
QY      324  ---GluThrAlaGluGluLeuAlaGluIleAlaSerGluAsnSerAspSerIle 342
           |||  ::|||  ::|||  ::|||  ::|||
Db      837  GTCCCCCAGCAGAGGAGAGTGTGAGTCTGTGCAGAGCCAGCCAGAGAGCCCGTG 896
           |||  ::|||  ::|||  ::|||
QY      343  TyrSerSerThrProGlu 348
           |||  |||  ::|||
Db      897  -----AGTGTGCCCCAG 908
           |||  |||  ::|||

```

Search completed: February 7, 2003, 00:01:23
 Job time : 86 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2003, 10:00:47 ; Search time 90 Seconds

(without alignments)
8584.721 Million cell updates/sec

Title: US-09-856-836-1

1630

Perfect score: 1 ttaccctaccgtagga.....atgccttaataaaaaa 1630

Sequence:

IDENTITY_NUC
Gap 10.0, Gape 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCRUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	805.4	49.4	1430	10	US-09-925-301-520
2	375.8	23.1	437	10	US-09-960-352-8644
3	243	14.9	313	10	US-09-815-343-740
4	243	14.9	313	10	US-09-998-598-1947
5	240.8	14.8	686	10	US-09-735-705-36
6	240.8	14.8	686	10	US-09-850-716A-36
7	240.8	14.8	686	10	US-09-897-778-36
8	237.2	14.6	311	9	US-09-920-455-140
9	213.6	13.1	1002	9	US-09-938-842A-510
10	170.4	10.5	361	10	US-09-998-598-2482
11	153.2	9.4	469	10	US-09-770-444-220
12	109.4	6.7	419	10	US-09-965-708-323
13	73.8	4.5	463	10	US-09-924-035A-278
14	65.2	4.0	262	10	US-09-923-876-3070
15	64.2	3.9	346	10	US-09-796-692-7155
16	53.6	3.3	727	9	US-09-771-035A-21
17	49.6	3.0	391	10	US-09-770-791-2
18	44.6	2.7	362	10	US-09-960-352-3338
19	44.6	2.7	436	10	US-09-960-352-2621

c 20	44.6	2.7	437	10	US-09-960-352-9500	Sequence 9500, Ap
c 21	43.6	2.7	473	10	US-09-969-373-1212	Sequence 1212, Ap
c 22	43	2.6	655	10	US-09-917-800A-971	Sequence 971, App
c 23	42.8	2.6	195	10	US-09-998-598-2374	Sequence 2374, Ap
c 24	41.8	2.6	469	10	US-09-864-864-96	Sequence 96, Appl
c 25	40.8	2.5	825	10	US-09-764-869-1980	Sequence 1980, Ap
c 26	40	2.5	271	10	US-09-878-574-6844	Sequence 6844, Ap
c 27	39.8	2.4	163	9	US-10-046-935-153	Sequence 153, App
c 28	39.8	2.4	163	9	US-09-878-178-153	Sequence 153, App
c 29	39.8	2.4	2643	10	US-09-925-300-407	Sequence 407, App
c 30	39.8	2.4	3075	9	US-10-125-237-29	Sequence 29, Appl
c 31	39.4	2.4	474	10	US-09-998-598-1749	Sequence 1749, Ap
c 32	39.4	2.4	3025	10	US-09-764-855-29	Sequence 29, Appl
c 33	38.2	2.3	404	10	US-09-960-352-11556	Sequence 11556, A
c 34	37.6	2.3	2703	9	US-09-938-842A-2472	Sequence 2472, Ap
c 35	37.6	2.3	7042	10	US-09-876-667-1	Sequence 1, Appl
c 36	37.6	2.3	7075	10	US-09-876-667-15	Sequence 15, Appl
c 37	37.4	2.3	370	10	US-09-867-701-4324	Sequence 4324, Ap
c 38	37.4	2.3	640681	10	US-09-790-988-1	Sequence 1, Appl
c 39	37	2.3	335	10	US-09-815-242-2817	Sequence 2817, Ap
c 40	37	2.3	773	10	US-09-770-445-903	Sequence 903, App
c 41	37	2.3	1215	10	US-09-939-980-63	Sequence 63, Appl
c 42	37	2.3	2406	10	US-09-815-242-3798	Sequence 3798, Ap
c 43	36.8	2.3	414	10	US-09-983-965-3051	Sequence 3051, Ap
c 44	36.4	2.2	1881	9	US-09-213-888-20	Sequence 20, Appl
c 45	36.4	2.2	1881	9	US-09-328-877A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-925-301-520
Sequence 520, Application US/09925301
Patient No. US200205308&1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 520
LENGTH: 1430
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (104)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature
LOCATION: (105)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature
LOCATION: (1428)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-520

Query Match 49.4%; Score 805.4; DB 10; Length 1430;
Best Local Similarity 82.0%; Pred. No. 2.2e-220;
Matches 1091; Conservative 0; Mismatches 201; Indels 38; Gaps 13;

QY 326 TTCACAGCAAGTGTGGATGCGGTCTCAGAGATGAATGATGACCCCTGCTCATAG 385
|||||
DB 39 TTCACAGCAAGTGTGGATGCGGTCTCAGAGATGAATGATGACCCCTGCTCATAG 98
|||||
QY 386 CACATGTCAGACAGTGTGATTTTCACAGATAGCAATTCACGATGATGACGAGGAGAC 445
|||||

Db	99	CACATNNNTCAAGACTGCTGAGATTTTCAACGACGAGATAGTATTTTCTTAAACCGGGGACAG	158
Qy	446	GATAAACGCTGCCCATATATGACTTGAACABAACCTGAAGACAAACCTTAAGAAATCACT	505
Db	159	GATAAACGCTTACGCATATATGACTTGAACAAACCTGAAGCAAACTTAAGAAATTAAT	218
Qy	506	GCCACACTTCTGGATTAATAAAGCCTCTGGGTCAGAGACGATTAACAGATCCCTTCA	565
Db	219	GCTACTACTTCTGGATTAATAAAGCCTCTGTGTGTCAGTGGAGATTAACAGATTCCTTCT	278
Qy	566	GCGAGTATTAACACTGTTCGGCTTGGGATCATGCGACAAATGACAGAACTGAATCTCTG	625
Db	279	GCTGATGACAAACACTGTTCGACTTTGGGATCATGCTACTATGACAGAACTGAATCTCTCA	338
Qy	626	AATTTTAATATGTCTGTTAGCAGCATGGAGTATATCTCCGAAGGAGATTTGGTTAT	685
Db	339	AATTTTAATATGTCTGTTAGTATGATATGATATATATCTCGAAGGAGATTTGGTTATA	398
Qy	686	ACTTATGACGATCTATTGCTTTTATATAGTGCAGTAACTCTGGAGCCAATTAATCTCTT	745
Db	399	ACTTATGACGATCTATTGCTTTTATATAGTGCAGTAACTTGGAGCCAATTAATCTCTT	458
Qy	746	GAGCTCCTGCAACATCAATTTCTGCTGTTNTTCAATCCAGAGAAAGATTTCTGTGG	805
Db	459	GAGCTCCTGCAACCATCAATTTCTGATCTCTTCACTGAGAAACAAATTTCTGTGCA	518
Qy	806	GGTGAGAAAGACTTTTAACTGTACAAGTATGATTTATACAGTGGAGAGATTAGATCC	865
Db	519	GCGCGGAAAGATTTTAACTTTATATAGTATATATATAGTGGAGAAAGATTAAGATCC	578
Qy	866	TACAAAGCTACTTTGTGCTCCCATCTACTGTGCAATTCAGTCCCTAGGGCAACTCTAT	925
Db	579	TACAAAGGACACTTTGTGCTCTATTCTACTGTGATGATTTATGCTGTAGTGAACCTCTAT	638
Qy	926	GCCAGCGGTTCTGAATATGGACATTTGAGATTTGTGCAAACTGTGTAGGAAAGACTTAT	985
Db	639	GCCAGTGTGTAGAAATATGSAACATTTGAGACATATGGCAAACTGTGTAGGAAACGTAAT	698
Qy	986	GGCCTGTGAAATGCTGNTTCTGTAGAGAAACAGCAGGGGAACTGGCAAAAGCCAAAGATC	1045
Db	699	GGCCTTTGGAAATGTGTCTCTCTGAAGAAATATGTGTAGTGGTGGCAAAAGCCAAAGATTT	758
Qy	1046	GGAATTCAGAAACAGACAGAGAGAGAGCTGGCAGAAATTTCTCTCAGAGAAATTCAGAT	1105
Db	759	GGTTTTCCAGAGCACACAGAAAGAGACT--AGAAATTTGCTTACAGAAATTCAGAT	815
Qy	1106	TCCATCTATTCATCAACTCTCTGAAGTTAAGGCTGAGCATACAGATGTGCTGCCGAAC	1165
Db	816	TGCATCTTCTCAGTCTCTGATGTTAAGGCTGAGCGTCAACATATGTGTGAGTTAG	875
Qy	1166	CAT--ATGTTCAATGAGCTAAACAAGACAGAACAGATCCGCTT--CAGAGTTACTGTCT	1222
Db	876	TATACACTACTATAAACAGACAGACAGAAAGAAAGCATCAGCCTTCCAGATTTACTGTCT	935
Qy	1223	GCCCTGAGGCAAAAGAGGCAAAAATATTGGGGCATATGAGTTAGCTTCAAGTCAGCAACA	1282
Db	936	GCTTAAGGCAAAACACAGTAATATATGAGAAATGAATTTACTCTCAGAGCTGAGCAACA	995
Qy	1283	GCT-----ACTCAGTGTGGCCCGTGAAGTGAAGATGGCTGAGTGTCTAGGTCGAGGGCAG	1337
Db	996	ACTTAACTAACCTTGGTGTACTCTGTAACTGAAAA---CTCAAGTGTCAAGATGAAGGACAG	1051
Qy	1338	AGGA-TTGTCCTCATAGTGCCTATACCTGTCTTT-----GGATGAAGAAAGCA	1387
Db	1052	TGAGATTAATCTCTATATAGTACATGTGCGCTTATATCTTTTAATGAATATATACAGACA	1111
Qy	1388	ACTTACATCTCCAT--TTACACCTAAATTTCTTTTAGCTTTATGT--TATGAGAA	1443
Db	1112	ACATCCAAATTTCTATTTATTAACAAATTAAGGCTTCTGTAGCTTTATATGTTAATGAGAA	1177
Qy	1444	GAATAATATATGGCAATTTTCTGACTTCCCTTAAGAAAGATGGCTTTTGTCTT	1503
Db	1172	GAATAATATATGGCTGATTTT-----TGTATCTTTAAAGCAAGATGCTTTTCTTTT	1227

Query Match	Best Local Similarity	23.1%	Score 375.8;	DB 10;	Length 437;
Matches 398;	Conservative	0;	Mismatches 38;	Indels	Gaps 0;
QY 495	AGGAATCATGAGTGCACACTTCTCGTATTAAAGGCTCTGTGGTGCATGACGATTAAC	554			
DB 1	AGGAATCATGAGTGCACACTTCTCGTATTAAAGGCTCTGTGGTGCATGACGATTAAC	60			
QY 555	AGATCCTTCAGGAGTATGATTAACCTGTGGCTCTGGAGCATGCCAATGACAGAG	614			
DB 61	AGATCCTTCAGGAGTATGATTAACCTGTGGCTCTGGAGCATGCCAATGACAGAG	120			
QY 615	TGAATCTCTGAATTTTAAATATGTCTGTAGCAGCATGAGATATATCTGGAAGGAGA	674			
DB 121	TGAATCTCTGAATTTTAAATATGTCTGTAGCAGCATGAGATATATCTGGAAGGAGA	180			
QY 675	TTTTGGTATTACTTATGAGCATCTATTGCTTTTCATAGTGCAGTAAGTCTGGACCAA	734			
DB 181	TTTTGGTATTACTTATGAGCATCTATTGCTTTTCATAGTGCAGTAAGTCTGGACCAA	240			
QY 735	TTAAATCTTTGAAGCTCTGCGACATCAATTCCTGCTCTNTTATCCAGAGAGAGT	794			
DB 241	TTAAATCTTTGAAGCTCTGCGACATCAATTCCTGCTCTNTTATCCAGAGAGAGT	300			
QY 795	TTTCGTGTTGGGGGAGAGACTTAAACTGATACAGTATGATTTAAACAGTGGAGAG	854			
DB 301	TTTCGTGTTGGGGGAGAGACTTAAACTGATACAGTATGATTTAAACAGTGGAGAG	360			
QY 855	AGTTAGAACTCTCAAGAGCTTGGTCCCATCTACGTGTGAGATTGACCTCGAAG	914			
DB 361	AATTAGAACTCTCAAGAGACTTGGTCCCATCTACGTGTGAGATTGACCTCGAAG	420			
QY 915	GGGAACCTATGCCAG 930				
DB 421	GAGAACTATGCCAG 436				

RESULT 3
 US-09-815-343-740
 : Sequence 740, Application US/09815343
 : Patent No. US20010055396A1
 : GENERAL INFORMATION:
 : APPLICANT: Meagher, Madeleine
 : APPLICANT: Xu, Jiangchun
 : APPLICANT: King, Gordon E.
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
 : FILE REFERENCE: 210121.504
 : CURRENT APPLICATION NUMBER: US/09/815.343
 : CURRENT FILING DATE: 2001-03-22
 : NUMBER OF SEQ ID NOS: 1556
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 740
 : LENGTH: 313
 : TYPE: DNA
 : ORGANISM: Homo sapien
 US-09-815-343-740

Query Match 14.9%; Score 243; DB 10; Length 313;
 Best Local Similarity 87.7%; Pred. No. 9.2e-60;
 Matches 277; Conservative 0; Mismatches 36; Indels 3; Gaps 1;

QY 821 AACTGTACAGTATGATTTATACAGTGAGAGAGATTAGATCTTACAAAGCTCATT 880
 DB 1 AACTTTATAGTATGATTTATATAGTGAGAGAAATAGATCTTACAAAGGACATTT 60
 QY 881 GTTCCATTGAGTGTGAGATTGAGTGTGAGAGAACTGTATGCCAGGCTTCTGAA 940
 DB 61 GTTCTATTGAGTGTGAGATTGAGTGTGAGAGAACTGTATGCCAGGCTTCTGAA 120
 QY 941 GATGGACATTGAGATTGTGGCAAACTGTGTGAGAAAGCCATATGCTGTGAAATGC 1000
 DB 121 GATGGACATTGAGATTGTGGCAAACTGTGTGAGAAAGCCATATGCTGTGAAATGT 180
 QY 1001 GTGNTTCTGAG 1060
 DB 181 GTGNTTCTGAG 240
 QY 1061 GCAG 1120
 DB 241 ACAG 297
 QY 1121 ACTCTGAGATTAGG 1136
 DB 298 GCTCCTGATGTTAAG 313

RESULT 4
 US-09-998-598-1947
 : Sequence 1947, Application US/09998598
 : Patent No. US20020150922A1
 : GENERAL INFORMATION:
 : APPLICANT: Stolk, John A.
 : APPLICANT: Xu, Jiangchun
 : APPLICANT: Chenault, Ruth A.
 : APPLICANT: Meagher, Madeleine Joy
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 : FILE REFERENCE: 210121.561
 : CURRENT APPLICATION NUMBER: US/09/998.598
 : CURRENT FILING DATE: 2001-11-16
 : NUMBER OF SEQ ID NOS: 2606
 : SOFTWARE: Corixa Invention Disclosure Database
 : SEQ ID NO 1947
 : LENGTH: 313
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-998-598-1947

Query Match 14.9%; Score 243; DB 10; Length 313;

Best Local Similarity 87.7%; Pred. No. 9.2e-60;
 Matches 277; Conservative 0; Mismatches 36; Indels 3; Gaps 1;
 QY 821 AACTGTACAGTATGATTTATACAGTGAGAGAGATTAGATCTTACAAAGCTCATT 880
 DB 1 AACTTTATAGTATGATTTATATAGTGAGAGAAATAGATCTTACAAAGGACATTT 60
 QY 881 GTTCCATTGAGTGTGAGATTGAGTGTGAGAGAACTGTATGCCAGGCTTCTGAA 940
 DB 61 GTTCTATTGAGTGTGAGATTGAGTGTGAGAGAACTGTATGCCAGGCTTCTGAA 120
 QY 941 GATGGACATTGAGATTGTGGCAAACTGTGTGAGAAAGCCATATGCTGTGAAATGC 1000
 DB 121 GATGGACATTGAGATTGTGGCAAACTGTGTGAGAAAGCCATATGCTGTGAAATGT 180
 QY 1001 GTGNTTCTGAG 1060
 DB 181 GTGNTTCTGAG 240
 QY 1061 GCAG 1120
 DB 241 ACAG 297
 QY 1121 ACTCTGAGATTAGG 1136
 DB 298 GCTCCTGATGTTAAG 313

RESULT 5
 US-09-735-705-36
 : Sequence 36, Application US/09735705
 : Patent No. US2002053239A1
 : GENERAL INFORMATION:
 : APPLICANT: Wang, Tonglong
 : APPLICANT: Fan, Liqun
 : APPLICANT: Kalos, Michael D.
 : APPLICANT: Bangur, Chaitanya S.
 : APPLICANT: Hosken, Nancy
 : APPLICANT: Fanger, Gary R.
 : APPLICANT: Li, Samuel X.
 : APPLICANT: Wang, Aljun
 : APPLICANT: Skelky, Yasir A.W.
 : APPLICANT: Henderson, Robert A.
 : APPLICANT: McNeill, Patricia D.
 : APPLICANT: Fanger, Neil
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 : FILE REFERENCE: 210121.455C14
 : CURRENT APPLICATION NUMBER: US/09/735.705
 : CURRENT FILING DATE: 2000-12-12
 : NUMBER OF SEQ ID NOS: 419
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 36
 : LENGTH: 686
 : TYPE: DNA
 : ORGANISM: Homo sapien
 : FEATURE:
 : NAME/KEY: misc.feature
 : LOCATION: (1)...(686)
 : OTHER INFORMATION: n = A,T,C or G
 US-09-735-705-36

Query Match 14.8%; Score 240.8; DB 10; Length 686;
 Best Local Similarity 79.9%; Pred. No. 6.4e-59;
 Matches 362; Conservative 0; Mismatches 83; Indels 8; Gaps 7;

QY 50 CGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 109
 DB 241 CTGAGCTGCGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 110 ACTGCTGCGGCGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 169
 DB 301 ACCTGCTGCGGCGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 359

OY 410 AACAGAGATGACATTTACTGTTAACTGGGGGACAGATAACTCTGGCATATATGAC 469
 Db 596 -CCCCAGATATATTTATTTTACCGGGGAGAGAT-AACTGTTTNCNTATTTTAA 653
 OY 470 TTGACAAACCTGAGAGACCTAAGGAATC 502
 Db 654 TTGAAACAA-CTMAAACAAACCTAAGGAATC 685

RESULT 8
 US-09-920-455-140

; Sequence 140: Application US/09920455
 ; Patent No. US2002016847A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tonglong
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.540
 ; CURRENT APPLICATION NUMBER: US/09/920.455
 ; NUMBER OF SEQ ID NOS: 275
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 140
 ; LENGTH: 311
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 115..199, 201, 235, 238
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-920-455-140

Query Match 14.6%; Score 237.2; DB 9; Length 311;
 Best Local Similarity 87.9%; Pred. No. 4.2e-58;
 Matches 254; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

OY 719 GTAAGTGTGAGACCAATTAATCTTTGAAGCTCTGGACATCAATTCGCTCTMT 778
 Db 23 GTAAGTGTGAGACCAATTAATCTTTGAAGCTCTGGACATCAATTCGCTCTMT 82
 OY 779 CATCAGAGAGAGATTTCTGTTGGGGTGGAGAGACTTTAACTTAAAGTATAT 838
 Db 83 CATCTGTGAGAGATTTCTGTTGGGGTGGAGAGACTTTAACTTAAAGTATAT 142
 OY 839 TATAAAGAGAGATTTCTGTTGGGGTGGAGAGACTTTAACTTAAAGTATAT 898
 Db 143 TATAAAGAGAGATTTCTGTTGGGGTGGAGAGACTTTAACTTAAAGTATAT 202
 OY 899 AGATTCACTCTGATGGGAACTATGCGGCTTCTGAAGATGGGCAATTTGAGATTG 958
 Db 203 AGATTCACTCTGATGGGAACTATGCGGCTTCTGAAGATGGGCAATTTGAGATTG 262
 OY 959 TGGCAAACTGTGTAGGAAAGACCTATGGGCTGTGAAATCGGNTTC 1007
 Db 263 TGGCAAACTGTGTAGGAAAGACCTATGGGCTGTGAAATCGGNTTC 311

RESULT 9
 US-09-938-842A-510

; Sequence 510: Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kieps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938.842A
 ; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 510
 ; LENGTH: 1002
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-510

Query Match 13.1%; Score 213.6; DB 9; Length 1002;
 Best Local Similarity 54.4%; Pred. No. 4.9e-51;
 Matches 473; Conservative 0; Mismatches 387; Indels 9; Gaps 2;

OY 102 CGCCGCTCACTTGTCTGGGGCCACAGCGGCGCTGTGATGATGCTTGCAGCGCATCA 161
 Db 8 CTCGGTTAGTGTGCGACGGGCAATCTGCGCGTGTGATGATGCTTGCAGCGCATCA 67
 OY 162 CGCCTTACGGCTACTTCTGATCAGCGCTTGCAGCAAGATGGCAAGCCATCTCCGAGG 221
 Db 68 CTCGAGATGGTCTTCTCTATTAAGCGCAGCAAGATGGCAAGCCATCTCCGAGG 127
 OY 222 GAGATACAGAGACTGTGATGGAATTTTGGCTTAAAGTGTGTTGGGGTGCMA 281
 Db 128 GAGAGACTGTGATGGAATTTTGGCTTAAAGTGTGTTGGGGTGCMA 187
 OY 282 CATGATTAAGATGACCAACCAAGCTGCGACAGCTGCAAGCTTCCACAGCAAGATAT 341
 Db 188 GCTTGTATTAAGATGACCAACCAAGCTGCGACAGCTTCCACAGCAAGATAT 247
 OY 342 GGGATGCGGCTCTGAGAGATGATGATGACCTGCTCTATGACATTTGCAAGACTG 401
 Db 248 GGAATGATGACAGAGATGATGATGACCTGCTCTATGACATTTGCAAGACTG 307
 OY 402 TGAATTCACAGAGATGATGATGATGACCTGCTCTATGACATTTGCAAGACTG 461
 Db 308 GTGCTTTCTGAGAGACACACCGTTTACTCACTGTGTGAAGAGCAAAATCTTCGGA 367
 OY 462 TATATGCTTGAACAACTGGAAGCAGACACTTAAGAAATCACTGTCGGA 521
 Db 368 TATGATTTGATGAG 427
 OY 522 TTAAG 581
 Db 428 TTAAG 487
 OY 582 TTGGGCTCTGGATCATGACCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637
 Db 488 GTGACATTTAGGATGAG 547
 OY 638 --TCTGTTACAG 695
 Db 548 CCCAGTTACTAGTGTGAG 607
 OY 696 GATCTATTTCTTTCAATGAG 755
 Db 608 CTAGTGTAAAGTTTGGAG 667
 OY 756 CGACATCAATTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 812
 Db 668 GCAATGTTGAATGGGATGTTGGAACCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 727
 OY 813 AAGACTTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 872
 Db 728 AAGATATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787
 OY 873 GTACATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 932
 Db 788 GTACATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 847

OY 933 GGTCTGAGATGGACATTTGATGTCG 961
| | | | | | | | | | | | | | | | | |
DB 848 GCTCAGAGACGGAACGGTCACATATGG 876

RESULT 10

US-09-998-598-2482
; Sequence 2482, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 2482
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-2482

Query Match 10.5%; Score 170.4; DB 10; Length 361;
Best Local Similarity 91.4%; Pred. No. 6.1e-39;
Matches 180; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 121 CCACAGCGCGCGGTGGATNTGGCTTCACAGCGCATCAGCCCTTACGCTACTTTC 180
| | | | | | | | | | | | | | | | | |
DB 1 CCACAGCGCACCGGTGGATNTGGCTTCACAGCGCATCAGCCCTTACGCTACTTTC 60
| | | | | | | | | | | | | | | | | |
OY 181 GATCAGCGCTTGGCAAGATGCGCAAGCCCATGCTCCGCCAGGAGATACAGAGACTGGAT 240
| | | | | | | | | | | | | | | | | |
DB 61 AATCAGCGCTTGGCAAGATGCGCAAGCCCATGCTCCGCCAGGAGATACAGAGACTGGAT 120
| | | | | | | | | | | | | | | | | |
OY 241 TGGAAATTTTGGTCATTAAGTGTCTGTTGGGGTGCACAACTTGAATTAAGATGCGAC 300
| | | | | | | | | | | | | | | | | |
DB 121 TGGAAATTTTGGTCATTAAGTGTCTGTTGGGGTGCACAACTTGAATTAAGATGCGAC 180
| | | | | | | | | | | | | | | | | |
OY 301 CAAAGCTGCGACAGCAG 317
| | | | | | | | | | | | | | | | | |
DB 181 CAAAGCAGCTACAGCAG 197
| | | | | | | | | | | | | | | | | |

RESULT 11

US-09-770-444-220
; Sequence 220, Application US/09770444
; Patent No. US20020023280A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: Ah, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Mathew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Moesner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maya
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2027 (PARA-016PRV)
; CURRENT APPLICATION NUMBER: US/09/770,444
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 220
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-444-220

Query Match 9.4%; Score 153.2; DB 10; Length 469;
Best Local Similarity 60.8%; Pred. No. 6.1e-34;
Matches 248; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

OY 120 GCCACAGCGCGCGGTGGATNTGGCTTCACAGCGCATCAGCCCTTACGCTACTTTC 179
| | | | | | | | | | | | | | | | | |
DB 5 GCGTCCGAGACCTGTTGGATTTGTTACAGTCCATCAGCTGATGATGTTCTTC 64
| | | | | | | | | | | | | | | | | |
OY 180 TGATCAGCGCTTGGCAAGATGCGCAAGCCCATGCTCCGCCAGGAGATACAGAGACTGGA 239
| | | | | | | | | | | | | | | | | |
DB 65 TCATCAGCGCAAGTAAAGATTCACACCAATGTTGAGAAATGGGAACTGAGATTGGA 124
| | | | | | | | | | | | | | | | | |
OY 240 TTGAACTTTTGGTCATTAAGTGTCTGTTGGGGTGCACAACTTGAATTAAGATGCGCA 299
| | | | | | | | | | | | | | | | | |
DB 125 TTGGTATTTTGAAGGTCATTAAGGTCGTCTGTGAGATTTTGTCTGATTAACATGCTT 184
| | | | | | | | | | | | | | | | | |
OY 300 CAAAGCTGCGACAGCAGCTTCACAGCCCAAGTAAAGTAAAGTAAAGTAAAGTAAAG 359
| | | | | | | | | | | | | | | | | |
DB 185 TAGTGAGCTTGCATGCTGATTTTTCAGGAGAGCTTTGGAGTCTTACCTGGG 244
| | | | | | | | | | | | | | | | | |
OY 360 ATGATTTGATGACCTTGGCTCATTAAGCAGATGTCAGAGATGAGATGATACAGAGATA 419
| | | | | | | | | | | | | | | | | |
DB 245 ATGCTTTTCTTTTGGACACAGCATTAATGTTGAGAGATGCGCTTCTCAGAGATA 304
| | | | | | | | | | | | | | | | | |
OY 420 GCAATTTACTGTAATGAGGAGACAGATTAAGTGTGCGCATTAATGATTAAGCAAAAC 479
| | | | | | | | | | | | | | | | | |
DB 305 CGAATTTCTTAATCAGAGAGATTTGAGAAATTTCTGCTGTTTTCATTAATGCT 364
| | | | | | | | | | | | | | | | | |
OY 480 CTGAGCAGAACCTTAAGCAATCAGTGCCACACTTCTGTAATTAANA 527
| | | | | | | | | | | | | | | | | |
DB 365 TGGATGACCTCCTACAGAAATGATTAATCTCTGTTCTATCAGAA 412
| | | | | | | | | | | | | | | | | |

RESULT 12

US-09-969-708-323/c
; Sequence 323, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 323
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-323

Query Match 6.7%; Score 109.4; DB 10; Length 419;

Best Local Similarity 70.0%; Pred. No. 2e-21;
Matches 299; Conservative 0; Mismatches 96; Indels 32; Gaps 10;

Qy 1226 TGAGGCAAGAGGCAAAATATTTGGCATATGATAGTCCAGCAGCAAGCT 1285
Db 418 TAAGGCAAGAGGCAAAATATTTGGCATATGATAGTCCAGCAGCAAGCT 359
Qy 1286 -----ACTGAGTGTGCTCCGCTGAGTGAAGGCTGAGTGTGAGTGCAGCAGGAG 1340
Db 358 AACTACTGTGTGCTCCGCTGAGTGAAGGCTGAGTGTGAGTGCAGCAGGAG 303
Qy 1341 A-TTGGTGCATATGATGCTGAGTGTGCTGAGTGTGAGTGCAGCAGGAG 1390
Db 302 ACTTATGCTCTTATATGATGCTGAGTGTGCTGAGTGTGAGTGCAGCAGGAG 243
Qy 1391 TACAAATCCCATTTTACACCTAAATTTCTTTAGCTTTTATGT---TATGAGAGAA 1446
Db 242 TCCAAATTTCTATTTATACAAATAGGCTTCTTATGCTTTATATATGAGAGAA 183
Qy 1447 AATATATTTGGCTATTTTCTGACTTCCCTTAAAGAGAAATGCTTTGCTCC 1506
Db 182 AACTATATTTGGCTATTTTCTGACTTCCCTTAAAGAGAAATGCTTTGCTCC 127
Qy 1507 T--AGTATGAAGAGAGAAATACATATTAAGTAAAGGCTTTGAGTGTGATGTA 1564
Db 126 TTAAGTGTGAAGAGAGAAATACATATTAAGTAAAGGCTTTGAGTGTGATGTA 69
Qy 1565 CAAGACTGCTTCAAGACAGTCA-TATTTTATTTATTAATTAATGCTTAAAT 1623
Db 68 TGTACACTGCTTCAAGACATTAATTTGTTTGTGTTCTTAATTAATGCTTAAAC 9
Qy 1624 AAAAAA 1630
Db 8 AAAAAA 2

RESULT 13
US-09-924-035A-278/c
Sequence 278, Application US/09924035A
Patent No. US20020142319A1
GENERAL INFORMATION:
APPLICANT: Glaxo, Jn
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2011US
CURRENT APPLICATION NUMBER: US/09/924,035A
CURRENT FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: US 60/148,784
PRIOR FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 900
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 278
LENGTH: 463
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(463)
OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-278

Query Match 4.5%; Score 73.8; DB 10; Length 463;
Best Local Similarity 64.7%; Pred. No. 3.3e-11;
Matches 108; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 802 TGGGGTGGAGAGCTTTAACTGTACAAATATTTTAAACAGTGGAGAGAGT 861
Db 463 TGGTGTGGAGAGATATTTGGTGTGACTTTTGAATTTCCACCGGAGAGAGAT 404
Qy 862 ATCTCAAAAGCTTGTCCATTCATCTGTGAGATGATGCTGAGAGAGAGT 921
Db 403 ATGCAACAANNACATCATGTCGGTTCACGTGTGAGATTTGACCAACAGGTGAGT 344

Qy 922 CTATGCGAGCGTCTTGAAGATGGACATTTGATTTGGCAAACTG 968
Db 343 TTATGCTTACAGGCTCTGAAGATGGTACATCATCAATTTGGCAAACTG 297

RESULT 14
US-09-923-876-3070
Sequence 3070, Application US/09923876
Patent No. US2002013958A1
GENERAL INFORMATION:
APPLICANT: Lalquid, Raghunath V.
APPLICANT: Kamigaki, Laura Y. (Ito)
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
FILE REFERENCE: PL-0012-1 CON
CURRENT APPLICATION NUMBER: US/09/923,876
CURRENT FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/298,329
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/085,331
PRIOR FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 6332
SOFTWARE: PERL Program
SEQ ID NO 3070
LENGTH: 262
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US2002013958A1 70016120H1
NAME/KEY: unsure
LOCATION: 45
OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-3070

Query Match 4.0%; Score 65.2; DB 10; Length 262;
Best Local Similarity 58.6%; Pred. No. 6.6e-09;
Matches 112; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 798 TTGTGGGGTGGAGAGCTTTAACTGTACAAATATTTTAAACAGTGGAGAGT 857
Db 7 TTGTAGCTGGAGAGAGATTTGGTGTGCTTCAATTTCTTCACTGCTGAAGAA 66
Qy 858 TAGAATCTCAAAAGCTTGTGCTCCATTCATCTGTGAGTGTGAGTGGG 917
Db 67 TACTGTGAAGAGGAGCATATGAGCGGTGCTGCTGCTGCTGCTGCTGCTG 126
Qy 918 AACTATATCCAGCGGTGTGAGATGGAGATTTGAGATTTGCGAACTGTGTAGAA 977
Db 127 AATCATACCATCGGCTGTCAGAAATGCGACCATTTAGATCTGGAGCTTACCCAGTTA 186
Qy 978 AGACCTATGCG 988
Db 187 ATGCTGATGAC 197

RESULT 15
US-09-796-692-7155
Sequence 7155, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
FILE REFERENCE: 2077,001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2003, 12:23:12 ; Search time 362 Seconds

(without alignments)
10140.213 Million cell updates/sec

Title: US-09-856-836-1
Perfect score: 1630
Sequence: 1 ttaccgcctccgctgagga.....atgcctctaataaaaaa 1630

Scoring table: IDENTITY-MUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_101002:*

- 1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1626	99.8	1630	21	AAAA9249
2	1011.4	62.0	1820	20	AAV59096
3	992	60.9	1847	19	AAV54162
4	917.4	56.3	2812	23	ABV22737
5	917.4	56.3	2812	23	ABV28565
6	812.8	49.9	15320	22	AAK78660
7	812.8	49.9	15320	22	AAK78660
8	805.4	49.4	1430	21	AAK78126
9	338	20.7	1331	23	ABL06979

10	321.8	19.7	522	24	ABO58251	Human colon cancer
11	245	15.0	1538	21	AAK46123	Arabidopsis thaliana
12	243.4	14.9	1540	21	AAK47955	Arabidopsis thaliana
13	243	14.9	313	23	AA558064	CDNA #740 encoding
14	240.8	14.8	686	20	AA224536	Human lung tumor a
15	240.8	14.8	686	21	AA65775	Human lung cancer-
16	240.8	14.8	686	24	ABK48994	Human lung tumour
17	237.2	14.6	311	24	ABK53945	Human head and nec
18	220.8	13.5	1753	21	AAK43745	Zea mays DNA fragm
19	213.6	13.1	1410	21	AAK47397	Arabidopsis thaliana
20	204.6	12.6	4891	23	ABL06978	Arabidopsis thaliana
21	201.6	12.4	1410	21	AAK38458	Arabidopsis thaliana
22	178.6	11.0	4179	23	ABL09294	Human genome fragm
23	170.8	10.5	223	15	AAO76451	Arabidopsis thaliana
24	153.2	9.4	469	24	ABL93455	Human full-length
25	128.8	7.9	1710	22	AAK94176	Prostate cancer re
26	109.4	6.7	419	19	AAV54156	Nucleotide sequenc
27	104.4	6.4	418	19	AAV54156	Human CDNA 5'-end
28	98.8	6.1	754	22	AAK91664	Human CDNA clone r
29	98.8	6.1	754	22	AAK93620	Arabidopsis thaliana
30	92	5.6	507	21	AAK45273	Arabidopsis thaliana
31	89.6	5.5	2901	23	ABL16514	Drosophila melanog
32	84.4	5.2	3165	23	ABL03412	Drosophila melanog
33	83	5.1	563	22	AAK92482	Drosophila melanog
34	80.4	4.9	1072	21	AAK43747	Human CDNA 3'-end
35	76.6	4.7	901	23	ABL16515	Drosophila melanog
36	70.8	4.3	2668	23	ABL18840	Human secreted pro
37	66.2	4.1	198	21	AAK25451	Human gene signatu
38	65	4.0	65	24	ABN55849	Human gene signatu
39	61	3.7	155	16	AAK19328	Aspergillus oryzae
40	60.6	3.7	344	22	AAI90810	Arabidopsis thaliana
41	59.4	3.6	681	21	AAK12868	Arabidopsis thaliana
42	56.8	3.5	1246	21	AAK3847	Arabidopsis thaliana
43	56.8	3.5	1253	21	AAK49546	Arabidopsis thaliana
44	56.8	3.5	1256	21	AAK33404	Arabidopsis thaliana
45	52	3.2	987	21	AAK43473	Arabidopsis thaliana

ALIGNMENTS

RESULT 1
ID AAA49249 standard; CDNA; 1630 BP.

25-SEP-2000 (first entry)

Serine-threonine kinase receptor associated protein (STRAP) DNA sequence.
Serine-threonine kinase receptor associated protein: STRAP; mouse;
transferring growth factor beta; TGF-beta; proliferation; tumor growth;
cytostatic; anti-arteriosclerotic; anti-diabetic; nephrotic; cancer;
atherosclerosis; diabetes; ss.

Mus sp.

Location/Qualifiers
86..1141

Key CDS
/tag= a
/product= "STRAP"
/note= "Serine-threonine kinase receptor associated protein"
/transl_except= (pos:143..145, aa:Xaa)
/note= "Xaa = unknown"
/transl_except= (pos:431..433, aa:Xaa)
/transl_except= (pos:776..778, aa:Xaa)
/transl_except= (pos:1004..1006, aa:Xaa)
/note= "Xaa = Leu, Ile"

WO200034310-A1.

OY 1561 TGTACAGAGCTGCTTACAGACGCTCATATTTTATCTATTAATAATGCTCTAA 1620
 DB 1561 TGTACAGAGCTGCTTACAGACGCTCATATTTTATCTATTAATAATGCTCTAA 1620
 OY 1621 AATAAAAAAA 1630
 DB 1621 AATAAAAAAA 1630

RESULT 2

AAV59096
 ID AAV59096 standard; DNA: 1820 BP.

AAV59096:

20-JAN-1999 (first entry)

TFAS protein coding sequence.

TFAS: transforming growth factor-beta receptor associated protein;

KW signal transduction; sepsis; toxic shock; autoimmune thyroiditis; asthma;

KW polymyositis; lupus erythematosus; osteoporosis; ulcerative colitis;

KW glomerulonephritis; osteoarthritis; vitreoretinopathy; therapy; human;

KW inflammatory disorder; cancer; ss.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Location/Qualifiers
 259..1311
 /*tag= a
 /product= TFAS
 US5834240-A.
 10-NOV-1998.
 28-MAR-1997: 9705-0828922.
 28-MAR-1997: 9705-0828922.
 (INCY-) INCYTE PHARM INC.
 Bandman O, Lal P;
 WPI: 1999-008710/01.
 P-PSDB: AAM73146.
 Transforming growth factor-beta receptor associated protein - useful
 for developing products for treating e.g. cancers, transplant
 nephropathy, bronchitis, hypertension, restenosis, sepsis,
 ulcerative colitis, asthma and wound healing
 Claim 4; Fig 1; 28pp; English.
 This sequence encodes the transforming growth factor-beta receptor
 associated protein (TFAS) of the invention. TFAS is involved in signal
 transduction. TFAS polypeptides or agonists can be used for treating
 disorders associated with decreased TFAS expression, e.g. sepsis, toxic
 shock, autoimmune thyroiditis, polymyositis, lupus erythematosus,
 osteoporosis, ulcerative colitis, asthma, glomerulonephritis,
 osteoarthritis, vitreoretinopathy, or wound healing. Antagonists or
 inhibitors of TFAS can be used for treating e.g. cancers and
 inflammatory conditions e.g. kidney fibrosis, chronic lung disease of prematurity,
 systemic sclerosis, pulmonary fibrosis, chronic lung disease of prematurity,
 infants, bronchitis, emphysema, scleroderma, rheumatoid arthritis,
 hypertension, or restenosis. The products can also be used for antibody
 production, detection, diagnosis and drug screening.
 Sequence 1820 BP: 516 A; 370 C; 411 G; 523 T; 0 other;

OY 81 CCGCATGSCATGAGGACAGACGCGCTCACTTCTGCGGCCACACGCGCGCGTGGTGG 140
 DB 254 CCGCATGSCATGAGGACAGACGCGCGCTCACTTCTGCGGCCACACGCGCGCGTGGTGG 140
 OY 141 ATNTGCGCTTCAGGCGATCAGCGCTTACGCGCTTCTGCGGCCACACGCGCGTGGTGG 200
 DB 314 ATNTGCGCTTCAGGCGATCAGCGCTTACGCGCTTCTGCGGCCACACGCGCGTGGTGG 200
 OY 201 GCAAGCCATGCTCCGCGAGGAGATACAGAGAGCTGATTTGGACATTTTGGTGCATTA 260
 DB 374 GCAAGCCATGCTCCGCGAGGAGATACAGAGAGCTGATTTGGACATTTTGGTGCATTA 260
 OY 261 AAGTGTCTTTGGGGGCAACATTTAGATGATCCACCAAGAGTGGCAGACAGCTG 320
 DB 434 AAGTGTCTTTGGGGGCAACATTTAGATGATCCACCAAGAGTGGCAGACAGCTG 320
 OY 321 CAGACTTCACAGCCAAAGATATGAGATGCGGTCTCAGAGATGATGATGACCTGCTC 380
 DB 494 CAGACTTCACAGCCAAAGATATGAGATGCGGTCTCAGAGATGATGATGACCTGCTC 380
 OY 381 ATTAAGACATTTGACAGACTGCTGATTTTACACAGAGATGCAATTAATTAATTAATTA 440
 DB 554 ATTAAGACATTTGACAGACTGCTGATTTTACACAGAGATGCAATTAATTAATTAATTA 440
 OY 441 GACAGATTAACCTGCTGCGCATATATGATGATGCAACCTGAGAGAGAGTAAAGAA 500
 DB 614 GACAGATTAACCTGCTGCGCATATATGATGATGCAACCTGAGAGAGAGTAAAGAA 500
 OY 501 TCACTGCGCACCTCTGCTGATTTAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
 DB 674 TCACTGCGCACCTCTGCTGATTTAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
 OY 561 TTTACGCGGATGATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620
 DB 734 TTTACGCGGATGATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620
 OY 621 CTCTGATTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
 DB 794 CTCTGATTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
 OY 681 TTTATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740
 DB 854 TTTATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740
 OY 741 CTTTGAAGCTCTGCGACATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
 DB 914 CTTTGAAGCTCTGCGACATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
 OY 914 CTTTGAAGCTCTGCGACATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
 DB 801 TTTGCGGCTGAGAGACTTTAACTGTAAGATGATGATTAAGAGAGAGAGAGAGAGAGAG 860
 OY 974 TTTGCGGCTGAGAGACTTTAACTGTAAGATGATGATTAAGAGAGAGAGAGAGAGAGAG 860
 DB 861 AATCCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
 DB 1034 AATCCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
 OY 921 TCTATGCGAGGCTTCTGAG 980
 DB 1094 TCTATGCGAGGCTTCTGAG 980
 OY 981 CCTATGCGCTTCTGAG 1040
 DB 1154 CCTATGCGCTTCTGAG 1040
 OY 1041 AGATGCGATTTCCAGAAAG 1100
 DB 1214 AGATGCGATTTCCAGAAAG 1100
 OY 1101 CAGATTCATCTATTCATCACTGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1160
 DB 1271 CAGATTCATCTATTCATCACTGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1160

QY 1161 GAACCAT--ATGTCATGACTAAACAGACAGACATCCGCTT-CAGACTAC 1217
 DB 1331 GTTAGATATCACTGACTAAACACAGACAGAGAAAGACAGCTCCAGAGTTAC 1390
 QY 1218 TGTCTCCCTGAGGCAAGAGGAGAAATATGCGGATATGATGTTAGCTCCAGTCCAC 1277
 DB 1391 TGTCTCTTAAGGCAAGACAGAGATATATGAGAAATATGATGTTAGCTCCAGTCCAC 1450
 QY 1278 GAACAGCT--ACTCAGTGTGCCCCGCTGAGTGAATAATGGCTGAGTCTGAGTCCAG 1332
 DB 1451 GACACACTACTACTACTGTTGCTTACCTGTAAGTGAATAA--CTCAGTGTGAGTGAAG 1506
 QY 1333 GCAGAGAGA-TTGCTCTACATAGTGCATAGCTGCTGTT-----GGATGAAA 1382
 DB 1507 GGAGGAGAGTATCTCTTATGATACAGTGCCTGTTATCTTTTAATGAAATATATACA 1566
 QY 1383 AGCAACTTCAATCTCCATTT--TTACACTTAAATTTCTTTAGCTGTTATGT--TAGT 1438
 DB 1567 AGCCAGATCCAAATTTCTATTTATTAACAATTAAGGTTCTTGAGCTGTTATGTAATATG 1626
 QY 1439 AAGAGAGAAATATATATGCGCTATTTTCTGACTTTCCTTAAAGAAATGCTTTTGTG 1498
 DB 1627 GAGAGAGAAATATATGCGCTATTTTCTGAT--CTTAAAGAGAAATGCGCTTTTCT 1682
 QY 1499 TCCCTGCT--AGTGAAGAGAGAGAAATATCATGATTAAGTAAAGGTTGATCTT 1556
 DB 1683 TTTTTCCTTCACTGTTGAAGAGAGAGAAATATCATGATTAAGTAAAGGTTGATCTT 1740
 QY 1557 TCATTTGACAGAGACTCTTCAGAGAGCTCA-TATTTTATGTTATCTAAATTAATGCTC 1615
 DB 1741 TCGTTTATTTGACACTGCTCTGAACATCTTAATTTGTTAGTTGCTTAATTAATGCTC 1800
 QY 1616 TCTTAATTAATTAATTA 1630
 DB 1801 TCTTAATTAATTAATTA 1815

RESULT 3
 AAV54162
 ID AAV54162 standard; CDNA; 1847 BP.
 XX
 AC AAV54162;
 DT 21-DEC-1998 (first entry)
 XX
 DE Nucleotide sequence 7 from human cell line SH-SY5Y.
 XX
 KW Human; SH-SY5Y; apoptosis; antibody; immunohistological staining;
 KW Inhibition; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 408..1103
 FT /*tag= a
 FT /product= "SH-SY5Y protein"
 XX
 PN W09839437-A1.
 XX
 PD 11-SEP-1998.
 XX
 PE 05-MAR-1998; 98WO-JP00905.
 XX
 PR 05-MAR-1997; 97JP-0050302.
 XX
 PA (KYOW) KYOMA HAKKO KOGYO KK.
 PI Sakaki Y;
 PI
 DR WPI; 1998-495844/42.
 DR P-PSDB; AAM54161.
 XX
 PT Novel apoptosis-related DNAs and proteins - for diagnosis,

PT preventing or treating diseases associated with apoptosis
 XX
 PS Claim 1; Pages 40-43; 70pp; Japanese.
 XX
 CC This is the nucleotide sequence of a nucleic acid of the human cell
 CC line SH-SY5Y, used in the method of the invention involving the use of
 CC novel apoptosis-related DNAs and proteins. The inventions can be used
 CC as diagnostic reagents for apoptosis e.g. (monoclonal) antibodies for
 CC the protein, as a reagent in immunohistological staining, as apoptosis
 CC inhibitors. It can also be used for treatment of apoptosis-related
 CC diseases.
 XX
 SQ Sequence 1847 BP; 513 A; 389 C; 422 G; 523 T; 0 other;
 QY
 Query Match 60.9%; Score 992; DB 19; Length 1847;
 Best Local Similarity 83.0%; Pred. No. 3,2e-276;
 Matches 1317; Conservative 0; Mismatches 229; Indels 40; Gaps 15;
 QY 66 TCCCTCCCTCCCTGTCGCCATGAGGAGGAGCGGCTGACTGCGGCCACA 125
 DB 275 TCGCTGGCTTCGCCCGCCATGCAATGAGACAGAGCGCCCTGCTGCGGCCACA 334
 QY 126 CCGGCGCCGCTGCTGATTTGCGCTTCAAGCGGATCAGGCTTACGGCTTCTGATCA 185
 DB 335 CGCAGCCGCTGCTGATTTGCGCTTCAAGCGGATCAGGCTTACGGCTTCTGATCA 394
 QY 186 GCGCTGCAAGATGCAAGCCATGCTCCGCA-GGAGATACAGAGAGACTGATTTGA 244
 DB 395 GCGCTGCAAGATGCAAGCCATGCTTACGCAAGGAGATGATTTGA 454
 QY 245 ACATTTTGGGTCAATTAAGGCTGCTTTGGGTCGCAATGATTAAGATGCAAGCCACA 304
 DB 455 ACATTTTGGGTCAATTAAGGCTGCTTTGGGTCGCAATGATTAAGATGCAAGCCACA 514
 QY 305 GCTGCGACAGAGCTGACAGCTTCAAGCCAAAGTATGCGCTGCTGAGAGATGAA 364
 DB 515 GCAGCTACAGAGAGCTGATTTTCAAGCCAAAGTATGCGCTGCTGAGAGATGAA 574
 QY 365 TTGATGACCTTGGCTCATTAAGCAGCTTGTCAAGAGCTGCTGATTTTCAAGAGATGCAAT 424
 DB 575 TTGATGACCTTGGCTCATTAAGCAGCTTGTCAAGAGCTGCTGATTTTCAAGAGATGCAAT 634
 QY 425 TACCTGNTAACTGGGGGAGAGATTAAGCTGCGCATATATGATGATTAAGATGCAAT 484
 DB 635 TATTTGTTAACTGGGGGAGAGATTAAGCTGCGCATATATGATGATTAAGATGCAAT 694
 QY 485 GCAGAACCTTAAGGAATGATGCGCAGACTTCTGTTATTAAGAGCTCTGTGTGAGT 544
 DB 695 GCAGAACCTTAAGGAATGATGCGCAGACTTCTGTTATTAAGAGCTCTGTGTGAGT 754
 QY 545 GACGATTAAGCAGATCTCTTCAAGGATGATTAAGCTGCGCTGCGGATGATGCAACA 604
 DB 755 GAGGATTAAGCAGATCTCTTCAAGGATGATTAAGCTGCGCTGCGGATGATGCAACA 814
 QY 605 ATGACAGAGTGAATCTCTGATTTTAAATGATGCTGTTAGCAGATGAGATATCTCT 664
 DB 815 ATGACAGAGTGAATCTCTGATTTTAAATGATGCTGTTAGCAGATGAGATATCTCT 874
 QY 665 GAAGGAGATTTGGTTATTAATGATGAGCAGATCTATTTGCTTTTCAATGATGAGAT 724
 DB 875 GAGGAGATTTGGTTATTAATGATGAGCAGATCTATTTGCTTTTCAATGATGAGAT 934
 QY 725 CCGGAGCCAAATTAATCTCTTGAAGCTCCGCGACCATCAATTAATCTCTGCTTCA 784
 DB 935 TTGGACCAATTAATCTCTTGAAGCTCCGCGACCATCAATTAATCTCTGCTTCA 994
 QY 785 GAGAGAGATTTCTGTTGCGGGTGAAGAGACTTTAACTGTACAGATGATTAAC 844
 DB 995 GAGAGAGATTTCTGTTGCGGGTGAAGAGACTTTAACTGTATGATGATTAAT 1054
 QY 845 AGTGAAGAGATTTAATCTCTCAAGAGCTGCTTTGCTTCCATGCTGAGATGCT 904
 DB 1055 AGTGAAGAGATTTAATCTCTCAAGAGCTGCTTTGCTTCCATGCTGAGATGCT 1114

OY 905 AGT-CCTGATGGGGAACCTCTATGCCAGCGCTTCTGAAGATGGGACATTGACATTGTGGCA 963
 DB 1115 ACTCCCTGATGGGGAACCTCTATGCCAGCGCTTCTGAAGATGGGACATTGACATTGTGGCA 1174
 OY 964 AACTGTGTAGGGAACCTATGGCTGGGAAATGGCTGTTCTCTGAGGAGGAGGAGGAGG 1023
 DB 1175 AACTGTGTAGGGAACCTATGGCTGGGAAATGGCTGTTCTCTGAGGAGGAGGAGGAGG 1234
 OY 1024 GGAAGTGGCAAGCCAAAGATCGATTTCAGAAACAGCAGAGGAGGAGGAGGAGGAGG 1083
 DB 1235 TGAGCTGGCAAGCCAAAGATCGATTTCAGAAACAGCAGAGGAGGAGGAGGAGGAGG 1291
 OY 1084 AATTGCTTCAGAGATTCAGATTCATCTATTCATCAACTCTCTGAAATTTAAGGCTGAGC 1143
 DB 1292 AATTGCTTCAGAGATTCAGATTCATCTATTCATCAACTCTCTGAAATTTAAGGCTGAGC 1351
 OY 1144 ATCAGAGCTGTGCTGCCAAACCAT--ATGTCATGGCTTAACAGAGGAGGAGGAGGAGG 1201
 DB 1352 GTCATCATATGTGTGCACTTGTATACACTGACTTAAACAGAGGAGGAGGAGGAGGAGG 1411
 OY 1202 CCGCCTT-CAGATTTACTGTCTGCTGAGGCAAGAGGAGGAGGAGGAGGAGGAGGAGG 1260
 DB 1412 CAGCTTCAGAGATTCAGATTCATCTATTCATCAACTCTCTGAAATTTAAGGCTGAGC 1471
 OY 1261 AGTAGCTCCAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1315
 DB 1472 AATTAGCTCCAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1527
 OY 1316 GAGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1373
 DB 1528 TCAAGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1587
 OY 1374 -----GGAAAGAAAGGCACTTCAATTCCTCAAT--TTACACCTAAATTTCTTTAG 1424
 DB 1588 TTTATGATTAAT 1647
 OY 1425 CTGTTATATGTT--TATGAAGAGAGAAATATATATATATATATATATATATATATAT 1481
 DB 1648 CTGTTATATGTT--TATGAAGAGAGAAATATATATATATATATATATATATATATAT 1703
 OY 1482 AGAAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1539
 DB 1704 AGAAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1763
 OY 1540 AACCGGTTGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1598
 DB 1764 AACTGCTTGTATTTCT--TCGTTCTATTTGACACCTGCTGAAACATCTAATGTTTAA 1821
 OY 1599 TATCTAAATTAAT 1624
 DB 1822 TGTCTAAATTAAT 1847
 RESULT 4
 ABV22737
 ID ABV22737 standard; cDNA; 2812 BP.
 AC ABV22737;
 XX 13-SBP-2002 (first entry)
 DE Human prostate expression marker cDNA 22728.
 XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX MO200160860-A2.
 XX PN 23-AUG-2001.
 XX PD

PF 20-FEB-2001; 2001MO-US05171.
 XX 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Endege WO, Monahan JE,
 PI WPI; 2001-662795/76.
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 PS Claim 1; Page 4000; 11750pp; English.
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV6213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 SQ Sequence 2812 BP; 778 A; 614 C; 601 G; 819 T; 0 other;
 Query Match 56.3%; Score 917.4; DB 23; Length 2812;
 Best Local Similarity 85.8%; Pred. No. 1.6e-234;
 Matches 1101; Conservative 0; Mismatches 165; Indels 17; Gaps 7;
 OY 39 CGGCTCTCGGGGCGCCCTCCCT 98
 DB 289 CGAGCAGCT 348
 OY 99 AGAGCGGCT 158
 DB 349 AGAGCGGCT 408
 OY 159 TCAGCGCTTACGCT 218
 DB 409 TCAGCGCTTACGCT 468
 OY 219 AGGAGATACAGAGATCTGATTTGGATTTGGATTTGGATTTGGATTTGGATTTGGATTT 278
 DB 469 AGGAGATACAGAGATCTGATTTGGATTTGGATTTGGATTTGGATTTGGATTTGGATTT 528
 OY 279 CAACATTAATTAAGATGCGCAAGAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 338
 DB 529 CAACATTAATTAAGATGCGCAAGAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588
 OY 339 TATGGATGCGGCTCTCAGAGATGAATTAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 398
 DB 589 TATGGATGCGGCTCTCAGAGATGAATTAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 648
 OY 399 CTGGGATTTTACACAGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 458
 DB 649 CTGGGATTTTACACAGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 708
 OY 459 GCATATATGACTGAACAACTGAGCAGAACTTAAGAAATGAGGAGGAGGAGGAGGAGGAGG 517

Db 649 CTGTGATTTACGAGGATAGTAATTTATTTGTAACCGGGGAGACAGTAATACGTGTAC 708
 Qy 459 GCATATATGACTTGAACAAACCTGAAGCAGACACTAAGAAATTCAGTGGCCACTTCT- 517
 Db 709 GCATATATGACTTGAACAAACCTGAAGCAGACACTAAGAAATTTAGTGTCTACTTCTG 768
 Qy 518 GGTATATTAAGAGGCTCTGTGTGCA-GTACGATTAACAG-ATCCTTCAGCGATGATA 575
 Db 769 GGTATTAAGAAAGCTCTGTGTGCAAGTGTGATTAACAGAAATCTCTCTGTATGACA 828
 Qy 576 AAACGTTCGGCTGGGATCATGCCACATGACAGAACTGTAATCTCTGATTTAATA 635
 Db 829 AAACGTTCGGCTGGGATCATGCCACATGACAGAACTGTAATCTCTGATTTAATA 888
 Qy 636 TGTCTGTTAGCAGCAGATGATATATTTCTGAGAGAGATTTGGTTATTTACTATGAC 695
 Db 889 TGTCTGTTAGTATGATGAAATATATTTCTGAGGAGAAATTTGGTTATTTACTATGAC 948
 Qy 696 GATCATATGCTTTTCATGTCAGTAAGTCTGAGCCCAATTAATCTTTAAGCTCTG 755
 Db 949 GATCATATGCTTTTCATGTCAGTAAGTCTGAGCCCAATTAATCTTTAAGCTCTG 1008
 Qy 756 GCACCATCAATTTCTGCTCTGTCATCCAGAGAAAGATTTCTGTCGGGCTGAGAG 815
 Db 1009 CAACCATCAATTTCTGCTCTGTCATCCAGAGAAATTTCTGTCAGGCGGTGAG 1068
 Qy 816 ACTTTAACTCTACAGATGATATTAATTAACAGTGAAGAGATTGAATCTTCAAGGTC 875
 Db 1069 ATTTTAACTTTATAGATGATATTAATTAAGTGAAGAAATTTGAATCTTCAAGGAC 1128
 Qy 876 ACTTTGTCCTCATCTGTCAGTGTGATTCCTGATGGGAGACTATGCGCGGTT 935
 Db 1129 ACTTTGTCCTCATCTGTCAGTGTGATTCCTGATGGGAGACTATGCGCGGTT 1188
 Qy 936 CTGAAGATGGCAGATTAAGTGTGGCAAACTGTGTAGAAAGCTATGCGCTGTGA 995
 Db 1189 CAGAAAGATGGCAATTAAGATGATGCAAACTGTGTAGAAAGCTATGCGCTGTGA 1248
 Qy 996 AATCGTGTTCCTGAGAGAGACGCGGAGACTGCGCAAAAGCCAAAGATGATTTCCAG 1055
 Db 1249 AATGTGTCTCTCTGAAAGAGATGATGAGCTGCGCAAAAGCCAAAGATGATTTCCAG 1308
 Qy 1056 AAACAGCAGAGAGAGCTGCGCAAGAAATTCCTGAGAAATTCGATTCATCATAT 1115
 Db 1309 AGACAAGAGAGAGACT-----AGAAATTCCTGAGAAATTCGATTCATCATAT 1362
 Qy 1116 CATCACTCTCTGAAATTAAGGCTGAGCATCAGACGTGTGCTCCGAAACCAT-ATGTT 1173
 Db 1363 CTTCAGCTCTCTGATGTTAAGGCTGAGCGCTCAATCATATGTTGCAATATACACTG 1422
 Qy 1174 CATGACTTAACAAGCAGAGACATCGGCTT-CAGAGTACTGTCTGCTGAGGCA 1232
 Db 1423 ACTTAAACAGCAGAGAGAAAGCAATCAGCTTCAAGATGATGCTGCTTAAAGGCA 1482
 Qy 1233 AAGAGGCGAGAAATTAATTTGGGCGATGATGATGCTCAGTGCAGACAGCT-----AC 1287
 Db 1483 GAAACGACAGTAATTAATGAGAAATGAATTAATGCTCCAGTGTGGAACAACACTAGCTAAC 1542
 Qy 1288 TCAGTGTGCGCGTGAAGTAAAA 1310
 Db 1543 TTGGTGTTCCTGTTAAGTAAAA 1565

RESULT 6
 ID AK78660/C
 AK78660 standard; DNA; 15320 BP.
 XX AK78660;
 DT 07-NOV-2001 (first entry)
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33472.

XX Human: immune: haematopoietic: immune/haematopoietic antigen; cancer;
 KW Cytostatic; gene therapy; vaccine; metastasis; ds.
 XX Homo sapiens.
 OS
 XX
 PN
 PD
 XX
 PF
 PE
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 02-MAR-2000; 2000US-0184564.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231245.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0231415.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234984.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483426/52.
 DR
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS
 PS Disclosure; SEQ ID NO 33472; 3071pp + Sequence Listing; English.
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 15320 BP; 4488 A; 2687 C; 3177 G; 4968 T; 0 other;
 Query Match 49.9%; Score 812.8; DB 22; Length 15320;
 Best Local Similarity 78.9%; Pred. No. 7.9e-224;
 Matches 1179; Conservative 0; Mismatches 270; Indels 46; Gaps 16;
 QY 163 GCCTTACGGCTACTTCTGATCAGCGCTTGCAAGATGGAAGCCCATGCTCCGCCAGGG 222
 DB 13046 GCTTATGATGATTTCTAATCATGCTTGCAAGCGCAACCTATGCTATGCGCAGG 12987
 QY 223 ACATACAGAGACGATGATGGAACATTTGGGTCAATAAAGTGGCTGTTGGGGTCAAC 282
 DB 12986 AGATACAGAGACGATGATGGAACATTTGGGTCAATAAAGTGGCTGTTGGGGTCAAC 12927
 QY 283 ATTGAATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
 DB 12926 ACTAATAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12871
 QY 343 GGATGCGGCTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 402
 DB 12870 GGATGCTATCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12811
 QY 403 GGATTCACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
 DB 12810 GGATTCATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12751
 QY 463 ATATGACTTGAACCAACCTGAAGCAGAACCTTAAGAAATCATGATGATGATGATGATGAT 522
 DB 12750 ATATGACTTGAACCAACCTGAAGCAGAACCTTAAGAAATCATGATGATGATGATGAT 12691
 QY 523 TAAAAAGCCTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 582
 DB 12690 TAAAAAGCCTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12631
 QY 583 TCGG-CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 641
 DB 12630 TAAGACTTTGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 12571
 QY 642 TTAGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 701
 DB 12571 TTAGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 701

Db 12570 CTAGTAGTATGGAATATATCTCTGAGGAGAGATTTGGTAAATACCTATAGCTGATCCA 12511
Qy 702 TTGCTTTTCATAGTCAGTACGAGTCTGAGGACCAATTAATCTTTGAGCTCTCGGAGCA 761
Db 12510 GTGCTTTTCATAGTCAGTACGAGTCTGAGGACCAATTAATCTTTGAGCTCTCGGAGCA 12451
Qy 762 TCAATTTCTGCTCTTTCATCCAGAGAGAGATTTCTGTTGGGCTGAGAGACTTGA 821
Db 12450 TTAACTGGAATCTCTGCTCTGAGGAGAGAGATTTCTGTTGGGCTGAGAGACTTGA 12391
Qy 822 AACCTGACAGATGATTAATACAGTGGAGAGATTTGAATCCATCAAGCTCACTTTG 881
Db 12390 AACT-TAATAGATATGATTAATAGTGGAGAGAGATTTGAATCCATCAAGAGCACTCTG 12332
Qy 882 GTCCCAATTCATGCTGTGAGATTCAGTCTGATGGGAGCACTGATCCAGCGCTTGTGAG 941
Db 12331 GTCTTATTCATGCTGTGAGATTTGATCCGATGAGAGACTGATCCAGCGCTTGTGAG 12272
Qy 942 ATGGCAATTTGATTTGGGCAAACTGTGTRGAGAAAGCTTATGGCTGTGGAATGCG 1001
Db 12271 ATGGAACACTGAGAC--TGGCAACTGTGTGAGAGAGAGAGATTTGGGCTTGGAAATGTG 12214
Qy 1002 TGTCTCTGAG 1061
Db 12213 TGTCTCTGAG 12154
Qy 1062 CAG 1121
Db 12153 GAG 12100
Qy 1122 CTCTGAG 1178
Db 12099 CTCTGAG 12040
Qy 1179 ACTAAACAG 1237
Db 12039 AACGAG 11980
Qy 1238 GGCAG 1296
Db 11979 AGCAATTAATTAATAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 11920
Qy 1297 CCGCTGAG 1356
Db 11919 CCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11861
Qy 1357 GCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1405
Db 11860 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11801
Qy 1406 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1462
Db 11800 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11741
Qy 1463 TTTTCTGAC-TTTTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1515
Db 11740 TTTTCTGAC-TTTTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11681
Qy 1516 AG 1575
Db 11680 TAAAG 11625
Qy 1576 TCAAG 1629
Db 11624 TCAAG 11570

RESULT 7

AAK78662/c

ID AAK78662 standard; DNA; 15332 BP.

AC AAK78662;

XX

DT 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33474.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
PN M0200157182-A2.
PD 09-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-0501354.
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229349.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.

PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234277
PR	21-SEP-2000	2000US-0234397
PR	25-SEP-2000	2000US-0234584
PR	25-SEP-2000	2000US-0234588
PR	26-SEP-2000	2000US-0235498
PR	27-SEP-2000	2000US-0235534
PR	27-SEP-2000	2000US-0235536
PR	29-SEP-2000	2000US-0236327
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	29-SEP-2000	2000US-0236370
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237040
PR	02-OCT-2000	2000US-0237049
PR	02-OCT-2000	2000US-0237053
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-02411785
PR	20-OCT-2000	2000US-02411787
PR	20-OCT-2000	2000US-02411788
PR	20-OCT-2000	2000US-02411789
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	20-OCT-2000	2000US-0241826
PR	01-NOV-2000	2000US-0244617
PR	01-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246534
PR	08-NOV-2000	2000US-0246535
PR	08-NOV-2000	2000US-0246536
PR	08-NOV-2000	2000US-0246537
PR	08-NOV-2000	2000US-0246538
PR	08-NOV-2000	2000US-0246539
PR	08-NOV-2000	2000US-0246540
PR	08-NOV-2000	2000US-0246541
PR	08-NOV-2000	2000US-0246542
PR	08-NOV-2000	2000US-0246543
PR	08-NOV-2000	2000US-0246544
PR	08-NOV-2000	2000US-0246545
PR	08-NOV-2000	2000US-0246546
PR	08-NOV-2000	2000US-0246547
PR	08-NOV-2000	2000US-0246548
PR	08-NOV-2000	2000US-0246549
PR	08-NOV-2000	2000US-0246550
PR	08-NOV-2000	2000US-0246551
PR	08-NOV-2000	2000US-0246552
PR	08-NOV-2000	2000US-0246553
PR	08-NOV-2000	2000US-0246554
PR	08-NOV-2000	2000US-0246555
PR	08-NOV-2000	2000US-0246556
PR	08-NOV-2000	2000US-0246557
PR	08-NOV-2000	2000US-0246558
PR	08-NOV-2000	2000US-0246559
PR	08-NOV-2000	2000US-0246560
PR	08-NOV-2000	2000US-0246561
PR	08-NOV-2000	2000US-0246562
PR	08-NOV-2000	2000US-0246563
PR	08-NOV-2000	2000US-0246564
PR	08-NOV-2000	2000US-0246565
PR	08-NOV-2000	2000US-0246566
PR	08-NOV-2000	2000US-0246567
PR	08-NOV-2000	2000US-0246568
PR	08-NOV-2000	2000US-0246569
PR	08-NOV-2000	2000US-0246570
PR	08-NOV-2000	2000US-0246571
PR	08-NOV-2000	2000US-0246572
PR	08-NOV-2000	2000US-0246573
PR	08-NOV-2000	2000US-0246574
PR	08-NOV-2000	2000US-0246575
PR	08-NOV-2000	2000US-0246576
PR	08-NOV-2000	2000US-0246577
PR	08-NOV-2000	2000US-0246578
PR	08-NOV-2000	2000US-0246579
PR	08-NOV-2000	2000US-0246580
PR	08-NOV-2000	2000US-0246581
PR	08-NOV-2000	2000US-0246582
PR	08-NOV-2000	2000US-0246583
PR	08-NOV-2000	2000US-0246584
PR	08-NOV-2000	2000US-0246585
PR	08-NOV-2000	2000US-0246586
PR	08-NOV-2000	2000US-0246587
PR	08-NOV-2000	2000US-0246588
PR	08-NOV-2000	2000US-0246589
PR	08-NOV-2000	2000US-024659

Query Match	Best Local Similarity	Score	DB 22:	Length	15332:
Matches 1179;	Conservative	0;	Mismatches 270;	Indels	Gaps
163	GGCTTACGGCTACTTTCGTGATCAGCGCTGTCGCAAGATGGCAACCCCATGCTCCGCGAGG	222			
13057	GTCCTTATGACTATTTCTTAATCATAGTCTTCGCAAGGCGAAMCCTATGCTATGCCAGGG	12998			
223	AGATACAGAGACTGGATTGGAACATTTTGGTCATTAAGTGCTGTTGGGTCGAAC	282			
12997	AGATACAGAGACTGGATTGGAACATTTTGGTCATTAAGTGCTGTTGGATGTCGAAC	12938			
283	ATTGAATTAAGAGTCCCAAGAGTGGGAGACAGAGCTGCAGATCTCACGCCCAAGATG	342			
12937	ACTGAATTAAGAGTCCCAAGAGTGGGAGACAGAGCTGCAGATCTCACGCCCAAGATG	12882			
343	GGATCGCGTCCAGAGATGAATTGATGACCCGCTCATTAAGCACATTTGTCAGACTGT	402			
12881	GGATCGCATATCCAGAGATGAATGATGACCCGCTCATTAAGCACATTTGTCAGACTGT	12822			
403	GGATTTCACACAGATAGCATTAACCTGCTGATGAGGAGGAGACAGATTAAGTGGCGGCAT	462			
12821	GGATTTCATCGAGATAGTAAATTTTGTGTTAAACGGGGACAGATTAAGTGGCGGCAT	12762			
463	ATATGACTTGAACCAACTGGAAGAGAACTTAAGGAATCAGTGGCCACACTTCTGGTAT	522			
12761	ATATGACTTGAACCAACTGGAAGAGAACTTAAGGAATCAGTGGCCACACTTCTGGTAT	12702			
523	TAAAGAGCTCTGTGTGTCAGTGAAGATTAACAGATCCTTTACCGGATGTAATAACTGT	582			
12701	TAAAGAGCTCTGTGTGTCAGTGAAGATTAACAGATCCTTTACCGGATGTAATAACTGT	12642			
583	TCGG-CTCTGGGATCATGCCCATGACAGAAAGGAATCTCTGAATTTTAAATTTGCTG	641			
12641	TAAAGAGCTCTGTGTGTCAGTGAAGATTAACAGATCCTTTACCGGATGTAATAACTGT	12582			

QY	642	TTGGCAGACATGGAGATATATTCGCGAAGGAGAGATTTTGGCTATTAATCTTATGACAGATCTA	701
Db	12581	CTAGTACTATGGAATATATTTCTGAGAGGAGAGATTTTGGTAATAACTTATGCTGTGATCCA	125222
QY	702	TTGCTTTTCATAGTGCAGTAAGTCTGGAGCAATTTAAATCTTTGAAAGCTCTGGACACA	761
Db	12521	GTGCTTTTGCACATACAGTAAGTATTTGGAACCAATTTAAATCCTTTCAAGCTCTGCAACCA	12462
QY	762	TCAAATCTGGCTGCTNTTCATCCAGANAAGAGATTTCTTTGCGGGGTGAGAAAGCTTTA	821
Db	12461	TTTACTCTGAAATCTTTCACACTTGAGAAAAGAAATTTCTTTGTCAGGTGGTGAAGATGTTA	12402
QY	822	AACTGTACAAGTATGATTTTAAACAGTGGAGAAAGATTTAAATCTTACCAAGATCTCATTTG	881
Db	12401	AACT-TATAAGTATGATGATTAATAGTGGAGAAAGAAATTAATTAATCCTTCAAGAGACTCTG	12343
QY	882	GTGCCATTCACGTGTGAAATTCAGTCCCTGATGGGGAAGCTATATGCCAGCGGTCTGAAAG	941
Db	12342	GTCTTATTCACGTGTGAAATTTGAATTCCTGATGGAGAAAGCTATATGCCAGGTGTTCTGAAAG	12283
QY	942	ATGGGACATTGAGATTTGGGCAACTGTGTGTGAGAAAGACATATGAGCTGTGGAAATATCG	1001
Db	12282	ATGGAACTGTGAGAC--TGGCAAACTGTGTGTGAGAAAGAAATATATGCTTGTGAAATATG	12225
QY	1002	TGTTTCTGAGAGAAAGACACGGGGGAAATCGGCAAAACCAAGATCGGATTTCCAGAAACAG	1061
Db	12224	TGCTTCTTAAGAAAGTACTGTGTGTGAGCTGTGGCAACCAAGATTTGTTTCCAAAGACAA	12165
QY	1062	CAGAGAGAGCTGGCAGAAAGAAATTTGCTTCAGAAATTCAGATTCATCTATTCATCA	1121
Db	12164	GAGAAAG-----CGTAGAAAGAAATTTGCTTCAGAAATTCAGATTCATCTATTCCTCAA	12111
QY	1122	CTCCGAATTTAAGGCTGAGCATCAGACGTGTGTGTGCGGAAACAT--ATGTTCAATGG	1178
Db	12110	CTCCCTCAAGTTAAGGCTGTGAAATTCATCAAGATGTGTGCAAGACAGATCAACTGTGACTAG	12051
QY	1179	ACTAAACAAGCAGAGACAAGATCCGCTT-CAGATTTACTGTGTGCTCTGAGCGCAAGAG	1237
Db	12050	AACGAGCAAGCAGAAAGAAAGATCATGCTTCCAGAGTTACTCTGCTTAAGGCAAAAC	11991
QY	1238	GGCAGAAAATATTTGGGGATATAGTTAGTTCGTCACAGTGCAGACAACAGCT-ACTCAGTGTG	1296
Db	11990	AGCAATTAATATATGAGGAATATGAAATTTAGTTCACATGTGTGAACAACTCAACATGCTGTTA	11931
QY	1297	CCCGTAGTGAAGATGCTGAGTGTGTGAGGTGAGCAGAGAGAGATTTGTGTCACATAGT	1356
Db	11930	CCCATTAAGTGAATCTCAAGAGTATACAGATG-AAGGGAGTGTGAGTTATCTCTCTGTAGT	11872
QY	1357	GGCATAGGCTG--CTGTTTGGATGAAGAAAGCCACTTACAA-----TCTCCATTTT	1405
Db	11871	ACAAATGGCTGTCACTTTTAAATGAATATGTACAAAGATCAATCAACTCTCTTATTTA	11812
QY	1406	ACACCTAAATTTCTTTTAAAGCTGTTATG--TTATGAAGAAGAAATATATTTGGCCTAT	1462
Db	11811	CAATTAAGAGTCTTGTGACGCTGTTTATGTCAATATGAGAGAAAGAAATATATTTGGCCTAT	11752
QY	1463	TTTTCTGAC--TTTTCCCTAAAGAAATGCGCTTTTCTCTGCTGAG-----TGATGA	1515
Db	11751	TTTGTGACTTTTCCCTTAAAGCAAGATGCTTTTGTGTGCTATGATGTTTTTGTGTTGGTTG	11692
QY	1516	AGAGAGGAGAAATACATGATAAAGTAACCGGTGTGATCTCTTTATTTTACAAAGACTGCT	1575
Db	11691	TAAAGAGGAGATACATGATAAAGTAACCGGTTTAATTTCTTCAATGTAC---ACTGGT	11636
QY	1576	TCAGACAGAGCTCA-TATTTTATGTTATCTTAAATTAATTTGGCTCAAAATTAATAA	1629
Db	11635	TCGTAACATCTAATATTTTATTTAGTGTCTTAATTAATAATGCTCTGTGAACAAAAA	11581

RESULT 8
AAC78126
ID AAC78126 standard; cDNA; 1430 BP

XX	AAC78126:
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Human cancer associated gene sequence SRQ ID NO:520.
XX	
KW	Human; cancer associated gene; cancer antigen; detection; cancer;
KW	diagnosis; cytoskeletal; proliferative; vulnery; immunomodulator;
KW	antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
KW	antiinflammatory; antihypoid; anti allergic; antibacterial; cardiant;
KW	dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW	vaccrotic; antisporitic; antiangiogenic; gene therapy; inflammation;
KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW	allergic reaction; graft versus host disease; organ rejection;
KM	haemostatic; thrombolytic; cardiovascular disorder; infection;
KM	neurological disease; drug screening; ss.
OS	Homo sapiens.
PN	WO200055350-A1.
XX	
PD	21-SEP-2000.
XX	
PF	08-MAR-2000; 2000WO-US05882.
PR	12-MAR-1999; 99US-0124270.
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Rosen CA, Ruben SM;
DR	WPI: 2000-587533/55.
DR	P-PSTDB: AAB43917.
XX	
PT	Novel isolated nucleic acids comprising sequences encoding peptides
PS	useful for treating or diagnosing e.g. cancer -
XX	
PS	Claim 1; Page 1050; 2352pp; English.
XX	
CC	AAC77607 to AAC78448 encode the human cancer associated proteins given
CC	in AAB83398 to AAB84239. The proteins can have activities based on the
CC	tissues and cells the genes are expressed in. Example of activities
CC	include: cytostatic; proliferative; vulnerary; immunomodulator;
CC	antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
CC	antiinflammatory; antihypoid; anti allergic; antibacterial; antiviral;
CC	dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC	nootropic; vaccrotic; antisporitic; antiangiogenic. The
CC	polynucleotides and polypeptides can be used for preventing, treating or
CC	ameliorating medical conditions and diagnosing pathological conditions.
CC	The present invention may be used to treat immune disorders by activating
CC	or inhibiting the proliferation, differentiation or mobilisation of
CC	immune cells, to treat disorders of haematopoietic cells, autoimmune
CC	disorders, allergic reactions, graft versus host disease and organ
CC	rejection, modulate haemostatic or thrombolytic activity, modulate
CC	inflammation, cancers, cardiovascular disorders, neurological disease and
CC	bacterial or viral infections. The peptides, nucleotides, antibodies,
CC	agonists and antagonists may be also be used in drug screens. AAC78449 to
CC	AAC78457 and AAB44240 represent sequences used in the exemplification of
CC	the present invention.
XX	
SQ	Sequence 1430 BP; 495 A; 220 C; 306 G; 406 T; 3 other;
	Query Match 49.4%; Score 805.4; DB 21; Length 1430;
	Best Local Similarity 82.0%; Pred. No. 2.8e-222;
	Matches 1091; Conservative 0; Mismatches 201; Indels 38; Gaps 13
OY	326 TTTCAGACCCTAATGATGGCGTCTCAGAGATGAATTGATGACCTGCCTCAAAAG 365
Db	39 TTTCAGACCCTAATGATGGCGTCTCAGAGATGAATTGATGACCTGCCTCAAAATA 98
OY	386 CACAATTGTCAAGACTGTGGATTTCACACAGATATGACCAATTACTCTGNTAACGTGGGGACAG 445

Db	99	CACATNNTCAGACTGTGGATTTCCAGGAGATGTATTTTGTCTAACCGGGGACAG	158
QY	446	GATAAACTGCTGGCATATATGACTTGAACAACACTGAACAGAACCTTRAGGAATTCAGT	505
Db	159	GATAAACTGTACGATATATGACTTGAACAACACTGAAGCAACCTTAAGGAATTTAGT	218
QY	506	GGCCACACTTGTGGATTAATAAAGGCTCTGGTGTGTCAGTACGATTAACAGATCCCTTTCA	565
Db	219	GCTCACTTCTTGGTATTAATAAAGGCTCTGTGTGTGTCAGTACGATTAACAGATCCCTTTCT	278
QY	566	GCAGATGATTAATAACTGTTCCGCTCTGGATCATGCCACATGACAGAGTGAATCTCTG	625
Db	279	GCTGATGACAAAACCTGTTCGACTTTGGGTCTAGTGTACTATACAGAACTGAAATCTCTA	338
QY	626	AATTTAATATGTCTGTATAGCAGCATGGAGTATATTCCTGAAGGAGAGATTTGGTTAT	685
Db	339	AATTTAATATGTCTGTATAGTATAGTATAGTAAATTTCTGAGGAGAGATTTGGTTAT	398
QY	686	ACTATATGACAGATCATTTGCTTTTATATAGTGAAGTAACTGTGAGCCAAATTAATCCTTT	745
Db	399	ACTATATGACAGATCATTTGCTTTTATATAGTGAAGTAACTGTGAGCCAAATTAATCCTTT	458
QY	746	GAAGTCTCTGGACATCATTTCTGCGTNTTTCATCCAGAGAGAAGATTTCTGTGGG	805
Db	459	GAAGTCTCTGGACATCATTTCTGCGTNTTTCATCCAGAGAGAAGATTTCTGTGGG	518
QY	806	GCTGAGAAAGATTTAAACTGTATCAAGTATATATATATACAGTGTGAGAAAGATATGATCC	865
Db	519	GCGGTGAAAGATTTTAACTTTAT	578
QY	866	TACAAAGTCTGACTTTGGTCCCATTCACCTGTGTAGATTCAGTCCATGSGGAACTCAT	925
Db	579	TACAAAGTCTGACTTTGGTCCCATTCACCTGTGTAGATTCAGTCCATGSGGAACTCAT	638
QY	926	GCCAGCGGTCTGGAAGATGAGGACATTTGAGATTTGGCAAACTGTGTAGGAAGAACTCAT	985
Db	639	GCCAGCGGTCTGGAAGATGAGGACATTTGAGATTTGGCAAACTGTGTAGGAAGAACTCAT	698
QY	986	GGCCTGTGAAATCGCTGNTTCTGTGAGGAACAGCGGGGAACTGGCAAAACCAAGATGC	1045
Db	699	GGCCTTTGAAATGTGTGCTCTGTGAAGAAATATGTGTAGTGTGGCAAAACCAAGATTT	758
QY	1046	GGATTTCCAGAAACACACAGAGGAAGAGCTGGCAAGAAATTCCTTCAGAGAAATTCAGAT	1105
Db	759	GGATTTCCAGAGCAACACAGAGAGAGCT--AGAGAAATTCCTTCAGAGAAATTCAGAT	815
QY	1106	TGCATATATTCATCAACTCCTGTGAAGTTAAGGCTTGAGCATCAGACGTGTCTGCCGAAC	1165
Db	816	TGCATATATTCATCAACTCCTGTGAAGTTAAGGCTTGAGCATCAGACGTGTCTGCCGAAC	875
QY	1166	CAT--ATGTTCATGACATTAACAAGCAGAGACAGCATCCGCCCT--CAGATTAATCTGCT	1222
Db	876	TATCAACACTGACTTAACAAGCAGAGACAGAAAGCAATCAGCTTCACAGATTAATCTGCT	935
QY	1223	GCGTGAAGCAAAAGGGCCAGAAATATTTGGGGCATATGAGTTAGCTCCAGTGCAGCAACA	1282
Db	936	GCGTGAAGCAAAAGGGCCAGAAATATTTGGGGCATATGAGTTAGCTCCAGTGCAGCAACA	995
QY	1283	GCT-----ACTCAGTGTGGCCCGTGAAGTGAAGATGGCTGAAGTGTGAAGTGCAGCAGG	1337
Db	996	ACTAATCAACTGTGTGTACTCTGTAAATGAAAA--CTCAAGTGTGAGATGAAGGGGAGG	1051
QY	1338	AGGA--TTGTGTCTACATATAGTGCCTTACCCTGTCTT-----GGAATGAAGAAAGCA	1387
Db	1052	TGGAGTATATCTCTTATATATAGCACTGGCTGTATATCTTTTATATATATATATATATATAT	1111
QY	1388	ACTTACATCTCATAT--TACACACTAATATCTTTTATGCTGTATATG--TATGAGAA	1443
Db	1112	ACATCCAAATTTCTATATTTTACAAATTAAGGTTCTGTAGCTGTATATATATATATATATAT	1171
QY	1444	GAATAATATATATGCTATATTTTCTGACTTCCCTTAAGAGAAATGCTTTTGTCTCT	1503

[illegible]

QY 128 CGCCCGTGTGATGCTTACGCGGATCAGCCCTTACGCTTACTTGTGATCAGC 187
 DB 396 CGACCCGCTGCTCACCCTGATTCACGACATCTGGACGCGGCTTCTCATCTCG 455
 QY 188 GCTTGCAAGATGAGCAAGCCCATGCTCCCGCAGGAGATACAGAGACTGATGGACA 247
 DB 456 GCTCTGCAAGATGAGCAAGCCCATGCTCCCGCAGGAGATACAGAGACTGATGGACA 515
 QY 248 TTTTGGCTCATTAAGTGTCTGTTGGGTGCAACATTTGAATGAAGTCCCAAGT 307
 DB 516 TTGAGAGGACACAAAGGCGCGCTGTGGAAAGCGGCTGAAAGCTGAAAGTCCAGCCTGGCC 575
 QY 308 GCGAAGAGAGCTGCAAGCTTACAGCCAAAGATGAGGCTGCTCAGAGATGAATG 367
 DB 576 GCTCTGCGAGCGCGGCTGCTCAGCCGCAAGGTGTGAATGCGGTGACCGGCGGCAAGTA 635
 QY 368 ATGACCTGCTGCTATGACATTTGTCAGAGCTTGTGATTTACAGAGATGCAATTAAC 427
 DB 636 CACAGCTTCCAGCACAAGACATCTGTGAAGAGCTGTGAGAGGATTCGAGAAC 695
 QY 428 CTGNTAATCTGGGGGAGAGATTAACCTGCTGCGCATATATGATGAAACCTGAAAGCA 487
 DB 636 ATAGTACTGTGAGCAAGAGAGTGTGCGGGCTTCAACCTGAGACAGCCGAGGCG 755
 QY 488 GAACCTAAGGAATCAGTGGCCACACTTCTGTATTAAGGCTCTGTGTCAGATGAC 547
 DB 756 CAGCCGCGAGATACGCGGACATACGCGGCGCATCAAGCGGCGCTCTTCTGCGCGCGC 815
 QY 548 GATAAAGATCTCTTACGCG---GATGATTAAGCTGTGCGCTGGGATCATGCCACA 604
 DB 816 GACAAGTCTATCTGCGGGCGGAGAGCAAGAGCGGCGCTCTGTGAGCCGATGACG 875
 QY 605 ATGACAGAGGAATCTCTCAATTTTAATATGCTGTGTAGACAGCATGAGATATATTCCT 664
 DB 876 GCGATCGAGGTGAGCGGCTGCAATTCAGAGCAATCCGAAACCTGAGAGATCTCCAGC 935
 QY 665 GAAAGAGATTTGTTGTTACTTATGAGAGCATATTTGCTTTTCAATAGAGCATAGT 724
 DB 936 GACAACCAATATGATGATATGCGACCGCTCTGATGATCTTCTGAGAGATGACAGC 995
 QY 725 CTGAGAGCAATTAATCTTGAAGCTCTGCGACCATCAATTCGTCTGTTTCATCA 784
 DB 996 CTGAGAGCAATTAATCTTGAAGCTCTGCGACCATCAATTCGTCTGTTTCATCA 1055
 QY 785 GAGAGAGATTTTCTTGTGGCGGTGAGAGAGATTTAACTTACATGATGATTAAC 844
 DB 1056 GATAGAGATGTTGTTGTTGTGGCGGCGAGAGACTTTAAGATGATGATGATGATC 1115
 QY 845 AGTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 904
 DB 1116 ACAGAGAAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1175
 QY 905 AGAGCTGATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 964
 DB 1176 AGCCCGGATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1235
 QY 965 AGTGGAG 1014
 DB 1236 ACCACCGTGGCAAG 1285

QY 392 GTCAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 451
 DB 3 GCGAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 62
 QY 452 CTGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 511
 DB 63 CTGTTAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
 QY 512 ACTTCTGATTAAG 571
 DB 122 TCTTCTGATTAAG 181
 QY 572 GATTAAG 631
 DB 182 GACAAAG 241
 QY 632 AATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 691
 DB 242 AATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
 QY 692 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 751
 DB 302 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 361
 QY 752 CCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 811

KW genetic analysis; diagnostic; antisense therapy; gene; ss.
 OS Homo sapiens.
 XX WO200229086-A2.
 XX 11-Apr-2002.
 XX 02-Oct-2001; 2001WO-US30732.
 XX 02-Oct-2000; 2000US-237271P.
 XX (FARB) BAYER CORP.
 XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA,
 XX Thiaalingam A, Lewis ME;
 XX WPI: 2002-426115/45.
 XX
 XX New isolated nucleic acid that is differentially expressed in cancer
 XX tissues useful for determining the presence of colon cancer in a cell
 XX or tissue type, and in antisense therapy
 XX
 XX Claim 1; Fig 1; 796pp; English.

AB06306 to AB060787 represent isolated nucleic acids (I) differentially
 expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
 encoded by the AB060776 to AB060787 nucleic acid sequences. (I) can be
 used in antisense therapy. An antibody immunoreactive with a polypeptide
 encoded by (I) is useful for detecting cancer in a patient sample, and
 for detecting the presence or absence of a polynucleotide encoded by a
 nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
 from (I) can be used for determining the presence of a nucleic acid which
 hybridizes to (I), and for determining the phenotype of cells in a sample
 of cells from a patient. (I) is useful for determining the presence of
 colon cancer in a cell or tissue type, for determining the presence or
 state of other type of cancer, in antisense therapy, to generate
 CC macroarrays on a solid surface, to identify a chromosome on which the
 CC corresponding gene resides, and in tissue profiling, forensics, genetic
 CC analysis, mapping and diagnostic applications. (I) can be used to raise
 CC antibodies, and to screen for peptide analogues and antagonists.

SO Sequence 522 BP; 149 A; 81 C; 114 G; 165 T; 13 other;
 Query Match 19.7%; Score 321.8; DB 24; Length 522;
 Best Local Similarity 84.5%; Pred. No. 1.5e-82;
 Matches 381; Conservative 0; Mismatches 65; Indels 5; Gaps 2;

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
D6	362	C	T	T	G	C	A	A	C	A	A	T	T	C	T	T	G	A	T	T	T	C	T	G	A	A	A	G	A	T	T	T	T	G	C	A	G	C	G	G	A	421																																																										
QY	812	G	A	G	A	C	T	T	T	A	A	C	T	G	T	A	C	A	T	A	G	T	A	T	A	842																																																																										
Db	422	N	A	---	---	T	T	T	A	C	T	T	T	A	A	G	T	N	G	A	T	T	T	A	448																																																																											

RESULT 11

ID AAC46123 standard; DNA; 1538 BP.

AC AAC46123;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48986

Hybridisation assay; genetic mapping; gene expression control;

KW metabolic pathway; promoter; termination sequence; ss.

05 Arabidopsis thaliana.

OS *Arabidopsis thaliana*.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR	5-FEB-1999	9905-0131825
PR	05-MAR-1999	9905-0133160
PR	09-MAR-1999	9905-0133548
PR	23-MAR-1999	9905-0125788
PR	29-MAR-1999	9905-0126268
PR	01-APR-1999	9905-0126785
PR	06-APR-1999	9905-0127462
PR	08-APR-1999	9905-0128234
PR	16-APR-1999	9905-0128714
PR	19-APR-1999	9905-0129845
PR	21-APR-1999	9905-0130077
PR	23-APR-1999	9905-0130449
PR	23-APR-1999	9905-0130892
PR	28-APR-1999	9905-0131048
PR	30-APR-1999	9905-0133044
PR	03-MAY-1999	9905-01332407
PR	04-MAY-1999	9905-0133484
PR	05-MAY-1999	9905-0133485
PR	06-MAY-1999	9905-0133486
PR	07-MAY-1999	9905-0133487
PR	11-MAY-1999	9905-0133863
PR	14-MAY-1999	9905-0134256
PR	14-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134219
PR	14-MAY-1999	9905-0134221
PR	18-MAY-1999	9905-0134370
PR	18-MAY-1999	9905-0134768
PR	19-MAY-1999	9905-0134941
PR	20-MAY-1999	9905-0135124
PR	21-MAY-1999	9905-0135353
PR	24-MAY-1999	9905-0135629
PR	25-MAY-1999	9905-0136021
PR	27-MAY-1999	9905-0136592
PR	28-MAY-1999	9905-0136782
PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137528
PR	04-JUN-1999	9905-0137502
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139119

PR	16-JUN-1999;	9905-01394552;
PR	6-JUN-1999;	9905-01394553;
PR	17-JUN-1999;	9905-01394592;
PR	18-JUN-1999;	9905-01394654;
PR	18-JUN-1999;	9905-01394655;
PR	18-JUN-1999;	9905-01394656;
PR	18-JUN-1999;	9905-01394657;
PR	18-JUN-1999;	9905-01394658;
PR	18-JUN-1999;	9905-01394659;
PR	18-JUN-1999;	9905-01394660;
PR	18-JUN-1999;	9905-01394661;
PR	18-JUN-1999;	9905-01394662;
PR	18-JUN-1999;	9905-01394663;
PR	18-JUN-1999;	9905-01394664;
PR	18-JUN-1999;	9905-01394665;
PR	24-JUN-1999;	9905-01406823;
PR	28-JUN-1999;	9905-01406824;
PR	30-JUN-1999;	9905-01406825;
PR	29-JUN-1999;	9905-01406826;
PR	01-JUL-1999;	9905-01412817;
PR	01-JUL-1999;	9905-01412818;
PR	02-JUL-1999;	9905-01420554;
PR	06-JUL-1999;	9905-01420555;
PR	08-JUL-1999;	9905-01420556;
PR	09-JUL-1999;	9905-01420557;
PR	12-JUL-1999;	9905-01420558;
PR	13-JUL-1999;	9905-01420559;
PR	14-JUL-1999;	9905-01420560;
PR	15-JUL-1999;	9905-01440005;
PR	16-JUL-1999;	9905-01440006;
PR	16-JUL-1999;	9905-01440007;
PR	19-JUL-1999;	9905-01440008;
PR	19-JUL-1999;	9905-01440009;
PR	19-JUL-1999;	9905-01440010;
PR	19-JUL-1999;	9905-01440011;
PR	19-JUL-1999;	9905-01440012;
PR	19-JUL-1999;	9905-01440013;
PR	20-JUL-1999;	9905-01440014;
PR	20-JUL-1999;	9905-01440015;
PR	20-JUL-1999;	9905-01440016;
PR	20-JUL-1999;	9905-01440017;
PR	21-JUL-1999;	9905-01440018;
PR	21-JUL-1999;	9905-01440019;
PR	21-JUL-1999;	9905-01450088;
PR	22-JUL-1999;	9905-01450089;
PR	22-JUL-1999;	9905-01450090;
PR	22-JUL-1999;	9905-01450091;
PR	22-JUL-1999;	9905-01450092;
PR	22-JUL-1999;	9905-01450093;
PR	23-JUL-1999;	9905-01450094;
PR	23-JUL-1999;	9905-01450095;
PR	23-JUL-1999;	9905-01450096;
PR	23-JUL-1999;	9905-01450097;
PR	26-JUL-1999;	9905-01452276;
PR	26-JUL-1999;	9905-01452277;
PR	27-JUL-1999;	9905-01452278;
PR	27-JUL-1999;	9905-01452279;
PR	27-JUL-1999;	9905-01452280;
PR	27-JUL-1999;	9905-01452281;
PR	27-JUL-1999;	9905-01452282;
PR	28-JUL-1999;	9905-01452283;
PR	02-AUG-1999;	9905-01453888;
PR	02-AUG-1999;	9905-01453889;
PR	02-AUG-1999;	9905-01453890;
PR	03-AUG-1999;	9905-01470388;
PR	04-AUG-1999;	9905-01470389;
PR	04-AUG-1999;	9905-01470390;
PR	05-AUG-1999;	9905-01471260;
PR	05-AUG-1999;	9905-01471261;
PR	06-AUG-1999;	9905-01471303;
PR	06-AUG-1999;	9905-01471416;
PR	09-AUG-1999;	9905-01474933;
PR	09-AUG-1999;	9905-01474934;
PR	10-AUG-1999;	9905-01481771;
PR	11-AUG-1999;	9905-01481819;

PN EPI033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128224.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134236.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134770.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136787.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140693.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 13-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149176.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

QY	520	TTTTAAAAAGGCTGCTGGTGACATGACGATTAACAGATCTCT-----TTACGGCGATGA	573
Db	799	TATCAGAACACTTAACATGGCTTCCACGGTATCGAACAAATTAACTTCTTGACACTGATAT	858
QY	574	TAAATCTGTGGGCTCTGGATGCATGCGACATATGACAGAGTAATAATCTCGAATTTTA	633
Db	859	TGCTGGTGTGAGGTTATGGGATGTGAGAGATGGCCAAATATGTCCAAACTGTGAAACCA	918
QY	634	TATGTCTGTAGACAGACATGAGATATATTCCTGGAAGAGAGATTTGGTTATTACTTATG	693
Db	919	GTCTCTGTACACAGTGTGTAAGTAGTCAAGATGAGCGGTATATACAACTGCCGATG	978
QY	694	AGGATCTATTTGGCTTTTCACTAGTGAAGTAATCTGTGAGCCAAATTAAATCTTTGAAGTCC	753
Db	979	GTCAACCTGTAAATTTTGGATGGCAATTCATTTTGGACTAGAGAAAAGTTACGACATGCC	1038
QY	754	TGCGACATCAATTCCTCGCTGCTTTCATCCAGAGAGG---AGTTTCCTGTGGGGGGTGC	810
Db	1039	CTGCAATTCGAATCTGATGCTGTAGGCCAAATCTGGCCAAATATGCTGGCTGGGG	1098
QY	811	AGAACTTTTAACTGTACAAAGTATGATTTAAACAGTGGAGAGATTAGATCCTACAA	870
Db	1099	AGAAATATGTGGGTGTGACTCTTTGTATTTCCACCGGAAAGAGATTTGATGTCAGAA	1158
QY	871	AGGTCACTTGGTCCCATTCATGCTGTGAGATTCAGTCTGTATGGGGAACTCTATGCCAG	930
Db	1159	GGGACATCATGTGCTGGTGTACTGTGTGAGATTTGGACCAACAGGTGTGATGTCT	1218
QY	931	CGGTCTGTGAAGATGGACATTTGAGATTTGGCCAAACTG	968
Db	1219	AGGCTGTGAAGATGTGTACAACTGAATTTTGGCAAAC	1256

RESULT 13
AAS58064
AAS58064 standard; cDNA: 313 bp.

XX Human: colon tumour protein.
KM CUBA #740 encoding portion of a human colon tumour protein.

XX	Homo sapiens.
OS	
XX	
XX	WO200173027-A2.
PN	
XX	
PD	04-OCT-2001.
XX	
PF	22-MAR-2001; 2001WO-US09246.
XX	
PR	24-MAR-2000; 2000US-191597P.
PR	04-MAY-2000; 2000US-202024P.
PR	05-MAY-2000; 2000US-202189P.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Meagher MJ, Xu J, King GE;
XX	
DR	WPI; 2001-611627/70.
XX	
PT	New colon tumour proteins and related nucleic acid, useful for
PT	treatment, prevention, diagnosis and monitoring of cancer -
XX	
PS	Claim 4; Page 174; 299pp; English.
XX	
CC	Th present invention relates to the isolation of novel cDNA sequences
CC	encoding for at least an immunogenic portion of human colon tumour
CC	proteins. The sequences of the invention are useful in pharmaceutical
CC	compositions and vaccines.

100

100

100

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

(without alignments)
10880.177 Million cell updates/sec

sequence: 1 ttatccgctaccgtaggga.....atgcctctaataataaaaaa 1630

gapop 10.0 , gapext 1.0

total number of hits satisfying chosen parameters: 4109280

Post-processing:	Minimum Match	0%
------------------	---------------	----

Maximum Match 100%
Listing first 45 summaries

Database :

1:	gb_ba:*
2:	gb_hrg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vl:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_ov:*
22:	em_or:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*
29:	em_vl:*
30:	em_hrg_hum:*
31:	em_hrg_inv:*
32:	em_hrg_other:*
33:	em_hrg_mus:*
34:	em_hrg_pin:*
35:	em_hrg_rod:*
36:	em_hrg_mam:*
37:	em_hrg_vrt:*
38:	em_sy:*
39:	em_hrgo_hum:*
40:	em_hrgo_mus:*
41:	em_hrgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

C	No.	Score	Query			DB	ID	Description
			Match	Length	DB			
C	1	1626	99.8	1630	10	AF096285	AF096285 Mus muscu	
C	2	1023.8	62.8	1867	9	HSM801659	AL136691 Homo sapi	
C	3	1020	62.6	1712	9	BC000162	BC000162 Homo sapi	
C	4	1020	62.6	1716	9	HA110025	AD010025 Homo sapi	
C	5	1018.4	62.0	1864	9	AB024327	AB024327 Homo sapi	
C	6	1011.4	62.0	1820	9	AR053395	AR053395 Sequence	
C	7	987.2	60.6	1563	9	AY049776	AY049776 Homo sapi	
C	8	934.2	57.3	15415	9	AC010878	AC010878 Homo sapi	
C	9	812.8	49.9	16680	9	AC010132	AC010132 Homo sapi	
C	10	812.8	49.9	187017	10	AC010101	AC010101 Homo sapi	
C	11	674.2	41.4	226520	2	AL6648802	AL6648802 Mouse DNA	
C	12	665	40.8	73038	2	AC020884.3	Continuation (4 of	
C	13	631.8	38.8	170393	10	AC109605	AC109605 Mus muscu	
C	14	358	33.0	1829	5	AF315726	AF315726 Cae	
C	15	338	20.7	1591	3	AY061000	AY061000 Drosophi	
C	16	285.6	17.5	413	9	HS056430	US6430 Human Hela	
C	17	243	14.9	313	6	AX261089	AX261089 Sequence	
C	18	240.8	14.8	686	6	AR144171	AR144171 Sequence	
C	19	240.8	14.8	686	6	AR176358	AR176358 Sequence	
C	20	240.8	14.6	686	6	AX365643	AX365643 Sequence	
C	21	237.2	14.6	311	6	AX393751	AX393751 Sequence	
C	22	213.6	13.1	1411	8	AY084604	AY084604 Arabidops	
C	23	212	13.0	1399	8	AY074392	AY074392 Arabidops	
C	24	204.6	12.6	31772	2	AC020465	AC020465 Drosophi	
C	25	204.6	12.6	188272	3	AC005639	AC005639 Drosophi	
C	26	204.6	12.6	21865	3	AC009257	AC009257 Drosophi	
C	27	178	10.9	182247	9	AC022073	AC022073 Homo sap	
C	28	170.8	10.5	223	6	A74365	A74365 Sequence	
C	29	170.8	10.5	223	6	A77344	A77344 Sequence	
C	30	134.2	8.2	60376	2	AC097388	AC097388 Rattus no	
C	31	134.2	8.2	169695	2	AC106673	AC106673 Rattus no	
C	32	125.8	7.7	169695	2	AC106673	AC106673 Rattus no	
C	33	125.2	7.7	51381	2	AF322456	AF322456 Homo sap	
C	34	109.4	6.7	419	6	AX337285	AX337285 Sequence	
C	35	89.6	5.5	89765	2	AC017970	AC017970 Drosophi	
C	36	89.6	5.5	188873	3	AC009800	AC009800 Drosophi	
C	37	89.6	5.5	196337	3	AC005894	AC005894 Drosophi	
C	38	89.6	5.5	242172	3	AE003832	AE003832 Drosophi	
C	39	83.6	5.1	2781	10	AF020683	AF020683 Rattus no	
C	40	82.2	5.0	34410	3	AC008242	AC008242 Leismani	
C	41	82.2	5.0	116000	2	AC129714	AC129714 Leismani	
C	42	75.2	4.6	98151	2	AC119404	AC119404 Magnapor	
C	43	66	4.0	10048	3	AY061931	AY061931 Drosophi	
C	44	64.6	4.0	81370	8	AB017071	AB017071 Arabidops	
C	45	64.6	4.0	11737	8	FB08	AC008016 Arabidops	

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AF096285	1630 bp	mRNA	linear	AF096285	AF096285.1	GI:4063382	Mus musculus.	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Datta,P.K., Chytil,A., Gorska,A.E. and Moses,H.L.
										Identification of STRAP, a novel WD domain protein in transforming

growth factor-beta signaling
J. Biol. Chem. 273 (52): 34671-34674 (1998)
MEDLINE 99074230
PUBMED 9856985
REFERENCE 2 (bases 1 to 1630)
AUTHORS Datta, P.K.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-1998) The Vanderbilt Cancer Center, Vanderbilt University School of Medicine, 649 Medical Research Building II, Nashville, TN 37232-6838, USA
Location/Qualifiers

FEATURES

CDS

1..1630
/organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="embryo"
86..1141
/note="WD domain protein; STRAP; TGF-beta receptor-associated protein"
/product="serine-threonine kinase receptor-associated protein"
/protein_id="AAC98300.1"
/db_xref="GI:4063383"
/translation="MAMROTPLTSGHTRPVYDLAFSGITPYGFLISACKGKPMR /OEDPWIGTIFGHKGVATLNDKATATADAFSTKWDVADSGDELMLAKHI VKTVDTPDSNITLITGDKLRIYDNPAPAPKISGHTSGIKKALMSDDKOLLS ADKTRVLDHATMTKSLNFMASVSMETIPREGIIVYIGRSIAFSAVSLPIK SFEAPATINSASLHPEKFLVAGDEDFKLYKTDYNGELESYKGFHGFHCVRSPP GELYASGSDGTLRIMQTYVVGKTYGLMKCVLPEDSGELAKPIGHPETAEBELAEI ASENSDIYSSTPEVKA"

BASE COUNT 455 a 349 c 406 g 420 t

ORIGIN

Query Match 99.8%; Score 1626; DB 10; Length 1630;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1626; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 TTTACCGCTACCGGTAGGACCGTGTAGTGGCGGACGCGGCTCTCCGGGCTCCCTCC 60
1 TTTACCGCTACCGGTAGGACCGTGTAGTGGCGGACGCGGCTCTCCGGGCTCCCTCC 60
61 CCTCCTCCCTCCCTCCCTGTGTCGCGCATGAGCCATAGGACGACGCGCTCACTTCTGGG 120
61 CCTCCTCCCTCCCTCCCTGTGTCGCGCATGAGCCATAGGACGACGCGCTCACTTCTGGG 120
121 CCACACGCGGCGCGTGTGTGATTTGGGCTTCAAGGCGATCAGCGCTTACGCGCTACTTCT 180
121 CCACACGCGGCGCGTGTGTGATTTGGGCTTCAAGGCGATCAGCGCTTACGCGCTACTTCT 180
121 CCACACGCGGCGCGTGTGTGATTTGGGCTTCAAGGCGATCAGCGCTTACGCGCTACTTCT 180
181 GATCAGCGCTTGCAGAAATGAGCAAGCCATCTCCGCGAGGAGATACAGAGACTGGAT 240
181 GATCAGCGCTTGCAGAAATGAGCAAGCCATCTCCGCGAGGAGATACAGAGACTGGAT 240
241 TGGAAATTTTTGGGTATTAAGTGTGCTTTGGGCTCAACATGATTAAGATGGCCAC 300
241 TGGAAATTTTTGGGTATTAAGTGTGCTTTGGGCTCAACATGATTAAGATGGCCAC 300
241 TGGAAATTTTTGGGTATTAAGTGTGCTTTGGGCTCAACATGATTAAGATGGCCAC 300
301 CAAAGCTGGAGAGAGAGTGCAGACTTCAAGCCAAAGTATGAGTGGCTCTCAGAGA 360
301 CAAAGCTGGAGAGAGTGCAGACTTCAAGCCAAAGTATGAGTGGCTCTCAGAGA 360
301 CAAAGCTGGAGAGAGTGCAGACTTCAAGCCAAAGTATGAGTGGCTCTCAGAGA 360
361 TGAATTTGATGACCTGGCTCATTAAGCTGCTTTGGGCTCAACATGATTAAGATGGCCAC 420
361 TGAATTTGATGACCTGGCTCATTAAGCTGCTTTGGGCTCAACATGATTAAGATGGCCAC 420
421 CAATTAAGCTGATTAAGTGGGAGCAGATTAAGTGTGATTTGATTTCAACAGAGATAG 480
421 CAATTAAGCTGATTAAGTGGGAGCAGATTAAGTGTGATTTGATTTCAACAGAGATAG 480
421 CAATTAAGCTGATTAAGTGGGAGCAGATTAAGTGTGATTTGATTTCAACAGAGATAG 480
481 TGAAGCAGAACTAAGGAATCATGTGGCACACTTCTGGTATTTAAAGAGGCTCTGTGGTG 540
481 TGAAGCAGAACTAAGGAATCATGTGGCACACTTCTGGTATTTAAAGAGGCTCTGTGGTG 540

541 CAGTACGATTAACAGATCCCTTACGAGGATGATTAACATGCTTGGCTCTGGATATGC 600
541 CAGTACGATTAACAGATCCCTTACGAGGATGATTAACATGCTTGGCTCTGGATATGC 600
601 CACATGACAGAGAGTGAATCTGTGAATTTTAAATATGTCTGTGTAGCAGATGAGATAT 660
601 CACATGACAGAGAGTGAATCTGTGAATTTTAAATATGTCTGTGTAGCAGATGAGATAT 660
661 TCCGTAAGAGAGATTTGTTGTTTACTTATGAGAGATCTATTCCTTTTCAATAGTGCAGT 720
661 TCCGTAAGAGAGATTTGTTGTTTACTTATGAGAGATCTATTCCTTTTCAATAGTGCAGT 720
721 AAGTGTGAGCCATTAATCCCTTGAAGCTCTGCGACATCAATTCCTGCTTNTTCA 780
721 AAGTGTGAGCCATTAATCCCTTGAAGCTCTGCGACATCAATTCCTGCTTNTTCA 780
781 TCCAGAGAGAGATTTTCTTTTGGGAGAGACACTTAAAGCTATCAGATATGATTA 840
781 TCCAGAGAGAGATTTTCTTTTGGGAGAGACACTTAAAGCTATCAGATATGATTA 840
841 TAAAGTGTGAGAGATTAAGATCTTACAAAGTCTCTTGGTCCATTCATCTGTGAG 900
841 TAAAGTGTGAGAGATTAAGATCTTACAAAGTCTCTTGGTCCATTCATCTGTGAG 900
901 ATTCACTCTGATGGGAACTCTATATGCCAGCGCTTCTGAATGAGGATTAAGATTTGTG 960
901 ATTCACTCTGATGGGAACTCTATATGCCAGCGCTTCTGAATGAGGATTAAGATTTGTG 960
961 GGAATCTGTGTGAGAGAAAGACCTATGAGCGCTGTGAATGCGGNTTCTGTGAGAGACAG 1020
961 GGAATCTGTGTGAGAGAAAGACCTATGAGCGCTGTGAATGCGGNTTCTGTGAGAGACAG 1020
1021 CCGGGAAGTGGCAAAAGCCAAAGATCGATTTCCAGAAACAGCAGAGAGAGAGCTGGAGA 1080
1021 CCGGGAAGTGGCAAAAGCCAAAGATCGATTTCCAGAAACAGCAGAGAGAGAGCTGGAGA 1080
1081 AGAATTTGCTTCAAGAAATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 1140
1081 AGAATTTGCTTCAAGAAATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 1140
1141 AGCATGAGAGAGTGTGCTGCGCAAAACATATGTTCAATGAGCTAAACAGCAGAGAGACA 1200
1141 AGCATGAGAGAGTGTGCTGCGCAAAACATATGTTCAATGAGCTAAACAGCAGAGAGACA 1200
1201 TCCGCTTCAAGATTTACTGTCTGCTGAGGCAAAAGAGGCAAAATATTTGGGCAATAG 1260
1201 TCCGCTTCAAGATTTACTGTCTGCTGAGGCAAAAGAGGCAAAATATTTGGGCAATAG 1260
1261 AGTTAGCTCAGTGCAGAGACACTACTGAGTGTGCGGTGAGTAAATGAGCTGAGTG 1320
1261 AGTTAGCTCAGTGCAGAGACACTACTGAGTGTGCGGTGAGTAAATGAGCTGAGTG 1320
1321 TCTGAGTGTGAGGAGAGAGATTTGCTCAGATGATGCTTACGCTGCTGTTGAGATTA 1380
1321 TCTGAGTGTGAGGAGAGAGATTTGCTCAGATGATGCTTACGCTGCTGTTGAGATTA 1380
1381 TCTGAGTGTGAGGAGAGAGATTTGCTCAGATGATGCTTACGCTGCTGTTGAGATTA 1440
1381 TCTGAGTGTGAGGAGAGAGATTTGCTCAGATGATGCTTACGCTGCTGTTGAGATTA 1440
1441 GAAAGCAATTTACATCTCAATTTTACCTTAATTTTCTTGTAGCTGTTATGATTA 1500
1441 GAAAGCAATTTACATCTCAATTTTACCTTAATTTTCTTGTAGCTGTTATGATTA 1500
1501 CTTGCTAGTGTGAGAGAGAGAGATTAAGTGAATGATTAAGTGAATGATTAAGTGAAT 1560
1501 CTTGCTAGTGTGAGAGAGAGAGATTAAGTGAATGATTAAGTGAATGATTAAGTGAAT 1560
1561 TGTCAAGAGAGTGTGAGAGAGAGAGATTAAGTGAATGATTAAGTGAATGATTAAGTGAAT 1620
1561 TGTCAAGAGAGTGTGAGAGAGAGAGATTAAGTGAATGATTAAGTGAATGATTAAGTGAAT 1620
1621 AATTAATAAATA 1630

Db 1294 TTGCTTCAGAGAAATTCAGATGATCTTCTTCTGAGTCTGATGTTAAGGCTGAGCGT 1353
 QY 1146 CAGACGTGCTGCTCCGGAACCAT-ATGTTTCATGAGCTAAACAAGACAGACATCC 1203
 1354 CAATCATATTTTGCAGTACTATACAACTGACATAAAACAAGCAAGCAAGAAAGCATCA 1413
 QY 1204 GCCTT-CAGAGTACTCTGTCTGAGGCAAGAGGCAAAAATNTTGGGGGATNTG 1262
 1414 GCCTTCAGAGTACTCTGCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1473
 QY 1263 TTACCTCAGAGGCAAGCAAGCT-ACACAGTGTGGCCGAGTGAAGAAATGGCTGA 1317
 1474 TTACCTCAGAGGCTGGAACACTACTACTGTGTGTACTGTAAAGTAAAA-CTC 1529
 QY 1318 GTGCTGAGTGCAGGAGGAGAGA-TTGCTGCTCAGATAGTCCATGCTGCTGTT- 1373
 1530 AAGTGTGAGTGAAGGAGGAGGAGTATCTCTTATAGTACATGCTGCTGTTT 1589
 QY 1374 -GGATGAAAGCAACTACATCTCAT-TTACCTTAATTTCTTTAGCT 1426
 1590 TAATGAATATATACAGCAATCCATTTCTATTTATTAATTAAGGCTTCTGAGCT 1649
 QY 1427 GTTATGTTATGAGAAAGAAATATATTTGCTATTTTCTGACTTCCCTTAAGAG 1486
 1650 GTTATGTTATGAGAAAGAAATATTTGCTATTTTCTGACTTCCCTTAAGAG 1709
 QY 1487 AATGCTTTTGTCTGCTTCCCT-AGTGAATGAGAGGAGGAAATCATGATTAAGTAA 1544
 1710 AATGCTTTTGTCTGCTTCCCTTAAAGAGAGGAGGAAATCATGATTAAGTAA 1769
 QY 1545 GTTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1603
 1770 GTTATGTTATGAGAAAGAAATATTTGCTATTTTCTGACTTCCCTTAAGAG 1827
 QY 1604 AATTAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1630
 1828 AATTAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1854

RESULT 3
 BC000162 1712 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, untr-interacting protein, clone MGC:5236
 DEFINITION IMAGE:2900954, mRNA, complete cds.
 ACCESSION BC000162
 VERSION BC000162.2 GI:14198026
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE Direct Submission
 AUTHORS Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On May 25, 2001 this sequence version replaced gi:12652818.
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalobdbcm.tmc.edu
 Villalob, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 3 Row: e Column: 15
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6807652.

FEATURES

source
 1. 1712
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /db_xref="locus:11171"
 /clone="MGC:5236 IMAGE:2900954"
 /tissue_type="Placenta, Chorioncarinoma"
 /clone_lib="NIH_MGC_10"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 146..1198
 /cdoon_start=1
 /product="untr-interacting protein"
 /protein_id="AAH00162.1"
 /db_xref="GI:12652819"
 /translation="MAMROTPLTSGHTRPVVDLAFSGITPYGYFLISACKDGMPLR
 OGDGDMWIGTFLGKGAIVGATLINDATAAADAADFAKVMADVSGDELMTLAKHI
 VRTVPTQDSNVLTLGDDKILRIYDLNKPPEAPEPEIGHTSGIKALMCSERQIIS
 ADKTVRLMDHAETPEVKSINFMNSVSMETPESEILVITYGHS IAFSAVSDPIK
 SEAPATVINSASLHPKEPEFLVAGDEFLKYADYNSGSELSYSGHSEPIHCVRSPD
 GELVAGSBDGTLRLMQTVYKATGLMKCVLPEDDSGELAPKIGFPETTELEELA
 SENSDCIPSPADPKA"

BASE COUNT

516 a 319 c 395 g 482 t

ORIGIN

Query Match 62.6%; Score 1020; DB 9; Length 1712;

Best Local Similarity 83.2%; Pred. No. 1.9e-262; Matches 1323; Conservative 0; Mismatches 229; Indels 38; Gaps 13;

QY 66 TCCCTCCCTCCCTGCTGCTCCGATGCGCATGAGCAGAGCGCGCTCACTGCTGCGGCACA 125
 Db 126 TCGGTGGCTTGGCGCGCGCATGAGCAGAGCGCGCTCACTGCTGCGGCACA 185
 QY 126 CGCGGCGCGGTGGTGGATTTGCGCTTACGCGCATCACGCTTACGCTTCTGATCA 185
 Db 186 CGCGACCGCGGTGGATTTGCGCTTACGCGCATCACGCTTACGCTTCTGATCA 245
 QY 186 GCGCTGCAAGAGTGGAGGAGCGCGCATGCGCGCATGAGCAGAGCTGATTTGGA 245
 Db 246 GCGCTGCAAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 305
 QY 246 CATTGTTGGTCAATTAAGGCTGCTGTTGGGCTGCAATTAAGGATTAAGGATTAAG 305
 Db 306 CATTGTTGGTCAATTAAGGCTGCTGTTGGGCTGCAATTAAGGATTAAGGATTAAG 365
 QY 306 CTGCGACAGAGCTGCGAGCTTACAGCGCAAGATGAGGATGCGGCTGCGAGAGTAA 365
 Db 366 CAGCTACAGAGCTGCGAGCTTACAGCGCAAGATGAGGATGCGGCTGCGAGAGTAA 425
 QY 366 TGATGACCTGCTGCTATTAAGCATTGTCAGAGCTGATTTGACACAGAGTAAATT 425
 Db 426 TGATGACCTGCTGCTATTAAGCATTGTCAGAGCTGATTTGACACAGAGTAAATT 485
 QY 426 ACCTGTTACTGCGGAGAGAGATTAAGCTGCGCATTAATTAAGTAAACCAACCTAAG 485
 Db 486 ATTGTTAACCGGGGAGAGATTAAGCTGCGCATTAATTAAGTAAACCAACCTAAG 545
 QY 486 CAGAACCTAAGAGATTAAGCTGCGCATTAATTAAGTAAACCAACCTAAG 545
 Db 546 CAGAACCTAAGAGATTAAGCTGCGCATTAATTAAGTAAACCAACCTAAG 605
 QY 546 ACATTAACAGATTCCTTTCAGCGGATTAAGCTGCGCATTAATTAAGTAAACCAAC 605
 Db 606 AGGATTAACAGATTCCTTTCAGCGGATTAAGCTGCGCATTAATTAAGTAAACCAAC 665
 QY 606 TGACGAGAGTAAATCTGAAATTTATGCTGTTAGCAGCAGTGGAGATTAATCTCG 665

Db	666	TGACAGAGGAATCTCTAAATTTTAAATATCTGTTAGTAGATGGAATATATTCCTG	725
Qy	666	AAGAGAGATTTGGTTATTTACTTATGAGACATCTATTGCTTTTCATAGTCAGTAGTC	725
Db	726	AGGGAAGATTTTGGTTATTAACCTTATGACGACATCTATTGCTTTTCATAGTCAGTAGT	785
Qy	726	TGGAGCAATTTAAATCCCTTGAAGCTCCCTCGACATCAATTTCTGCTCTTTTCATCAG	785
Db	786	TGAGACCAATTTAAATCCTTTGAAAGCTCTGACATCTATTGCTTTTCATAGTCAGTAGT	785
Qy	786	AGAAGAGTTTCTGTTGGGGGTGGAGAGACTTTAACTTCACTATGATATATTAACA	845
Db	846	AGAAAGATTTCTGTGTGACGGCTGAAGATTTTAAACCTTTAAATGATATATTAATA	905
Qy	846	GTGGAGAAAGTTAGATATCTACAAAGCTCACTTTGGTCCATTCAGTGTGAGATTA	905
Db	906	GTGGAGAAATTTAAATCTACAAAGGACACTTTGGTCCATTCAGTGTGAGATTA	965
Qy	906	GTTCATATGGGAGACATATGCGACGGCTCTGAAATGTGGACATTTGACATTTGGCAAA	965
Db	966	GTCTATATGGAACATCTATGCGACGGCTCTGAAATGTGGACATTTGACATTTGGCAAA	1021
Qy	966	CTGTGTAGGAAAGACCTATGGCCTGTGAAATCTGTGNTTCCGTGAGAGAACAGGGGG	1022
Db	1026	CTGTGTAGGAAAGACCTATGGCCTTTGGAAATGTGTCTCTCGAAGAAATAGTGTG	1081
Qy	1026	AACGTGCAAAAGCCAAAGATTCGATTTCCAGAGACACAGAAAGAGAGCT---AGAGGAA	1081
Db	1086	AACGTGCAAAAGCCAAAGATTCGATTTCCAGAGACACAGAAAGAGAGCT---AGAGGAA	1141
Qy	1086	TTTGCTTCAGAGAAATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCAT	1141
Db	1143	TTTGCTTCAGAGAAATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCAT	1201
Qy	1146	CAGAGTGTGCTGCCGAACCAT---ATGTTCAATGACATAACAGCAGACAGACATCC	1203
Db	1203	CAATATATGTTGCAAGTAGATTCACACTGATTAACAAAGCAGCAGAGAAACATCA	1263
Qy	1204	GCCCTT-CAGAGTATCTGCTGCTCGAGCAAGAGGGGAGGAAATTTTGGGCATATGAG	1263
Db	1263	GCCCTTCCAGATTTACTGCTGCTTAAAGGCAAGAACAGCATTAATATGAGGAAATGAA	1322
Qy	1263	TTAGCTCCAGTGCACGAACAGCT---ACTCAGTGTGCCGTGAGTGAATATGCTGCA	1317
Db	1333	TTAGCTCCAGTGCAGGAACTAATCACTTGTGTACGTGAATGAGGAA---CTC	1378
Qy	1318	GTGCTGAGGCGAGGAGAGAGA-TTGTGCTACATATGCCATATGCTGCTGTT---	1373
Db	1379	AAGTGTCAATGAGGAGGAGGATTCCTTTTAACTAAGGCGCTGTATCTTTT	1438
Qy	1374	-----GGAATGAAAGCCAACTTACAACTCCAT-TTACACTTAATTTCTTTTACT	1426
Db	1439	TAAATGATATATACAAACCAATCCAAATTTCTATATATACATTAAGGTTCTGTACT	1498
Qy	1427	GTATATGT---TATGAGAAAGAAATATATATGCGCTATTTTTCAGCTTTCCCTTAAG	1483
Db	1499	GTATATGTATATATGAGAAAGAACTAATATGCGTGATTTTTCGTAT---CTTAAG	1554
Qy	1484	AAGATGCTTTTGTCTGCTGCT-AGTATGAGAGAGAGGAAATATATATATATATAT	1541
Db	1555	CAGATGCTTTTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	1614
Qy	1542	CCGATTTGATCTTTTCATTTGACAAAGACTGCTTCAGAGACTCA-TATTTAGTTA	1600
Db	1615	CTGTTTGTATTT-TCGTTCAATGTACACTGCTCTGACATCATATGTTTTTATTTG	1672
Qy	1601	TCTTAATTAATATGCTCTTAATTTTAAAAA	1630
Db	1673	TCTTAATTAATATGCTCTTAATTTTAAAAA	1702

LOCUS	DEFINITION	1716 bp	mRNA	linear	PRI 12-MAR-1999
ACCESSION	HAJ010025				
VERSION	AJ010025				
KEYWORDS	AJ010025.1 GI:4468217				
SOURCE	RNA-binding protein; untr-interacting protein; unrip gene.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.				
AUTHORS	1 (bases 1 to 1716)				
TITLE	Hunt,S.L.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (06-AUG-1998) Hunt S.L., University of Cambridge, Department of Biochemistry, 80 Tennis Court Road, Old Addenbrooke's				
AUTHORS	Site, Cambridge, CB2 1GA, U.K.				
TITLE	2 (bases 1 to 1716)				
REFERENCE	Hunt,S.L., Huan,J.J., Totty,N. and Jackson,R.J.				
AUTHORS	unt, a cellular cytoplasmic RNA-binding protein with five				
TITLE	cold-shock domains, is required for internal initiation of				
JOURNAL	translation of human rhinovirus RNA				
MEDLINE	Genes Dev. 13 (4), 437-448 (1999)				
PUBMED	99158740				
FEATURES	10049359				
SOURCE	location/Qualifiers				
	1..1716				

BASE COUNT	ORIGIN
519 a	320 c
	395 g
	482 t

Query Match	62.6%;	Score 1020;	DB 9;	Length 1716;
Best Local Similarity	83.2%;	Pred. No. 1.9e-262;		
Matches 1323;	Conservative	0;	Mismatches 229;	Indels 38; Gaps 13

QY	66	TCCCTCCCTCCCTG	ATCCGCGCATGAGGACAGCCGGCTCACTTGGTGGGCCA	12
Db	128	TCCGTGGCTTCCGCCGCCG	CGCATGAGCAATGACAGACGGCGCTACCTCTCTGGCCACA	18
QY	126	CGCGGCCCGTGGT	GTGATNTGCGCTTCAAGCGGCATCACGCCCTTAGGGCTACTTTCATCA	18
Db	188	CGCGACCCGTGGT	ATTATTTGGCGCTTCAAGTGCATCACGCCCTTAGGGTATTCTTAATCA	24
QY	186	GGCCTTGCAAAGAT	GTGCACAGCCCATCGCTCCGCCACGGAGATACAGAGACCTGGATTGGA	24
Db	248	GGCCTTGCAAAGAT	GTGTAACCTTAGCTACGCCAGGAGATACAGGAGACTGGATTGGA	30
QY	246	CATTTTGGGCTAT	ATAAGGTCTGTTTGGGGTGCAACATTGAATAAGATGCCACCAAG	30
Db	308	CATTTTGGGCTAT	ATAAGGTCTGTTTGGGGTGCAACACTGATGAATGAAGTCACCAAG	36
QY	306	CTGCGACAGACAGCT	CACACTTCAACGCCAAAGATATGGGTGAGGCTCAGAGATGTAAT	36
Db	368	CAGCTGCACACAGCT	CACATTTCACACACAAATGTGGATGCTGTCTCAGAGATGTAAT	42

Oy	366	TGATGACCTGGCTATATAAGCAATATTCAAAGAGCTGGATTTTCACACAGGATTAACAT	425
Oy	367	TT	426
Db	428	TGATGACCTGGCTATATAAGCAATATTCAAAGAGCTGGATTTTCACACAGGATTAACAT	487
Oy	426	ACCTGATTAATCTGGGGACAGGATAAAGCTGCTGGCATATATGACTTTGAACAAACCTGAAG	485
Db	488	ATTTGTTTAACGGGGGACAGGATAAAGCTGATAGCAATATATGACTTTGAACAAACCTGAAG	547
Oy	486	CACAGCCTTAAGAAATTCAGTGGCCACACTTCTGATATAAAAGGCTGTGGTGACAG	545
Db	548	CACAGCCTTAAGAAATTAATGTTGTCATCACTTCTGATATAAAAGGCTGTGGTGACAG	607
Oy	546	ACGATTAACAGATCCCTTTCAGGGATGATAAAGCTGTGGCTCTGGATCATGCCACA	605
Db	608	AGGATTAACAGATCTTCTTCTGCTGATGACAAAGCTGTGGCTGGATCATGCTACGA	667
Oy	606	TGACAGAGTGAATCTCGAATTTTAATATGTCTGTTAACAGATAGAGATATATCCG	665
Db	668	TGACAGAGTGAATCTCTAAATTTTAAATATGTCTGTAGTAGATGAATATATTCG	727
Oy	666	AAGGAGAGATTTTGGTTTACTTATAGACATCATATGCTTTTCAATGACAGTAAGTC	725
Db	728	AGGAGAGATTTTGGTTTATTAATCTATAGACATCATATGCTTTTCAATGACAGTAAGTC	787
Oy	726	TGAGCCAAATTAATCTTTTGAAGCTCCTGGACCATCATTTCTGGCTCTNTTCAATCCAG	785
Db	788	TGAGCCAAATTAATCTTTTGAAGCTCCTGGACCATCATATGCTTTTCAATCCAG	847
Oy	786	AGAGAGAGTTTCTGTTTGGGGGTGGAGAAAGCTTAAACCTGACAAATATGATTTATACA	845
Db	848	AGAGAGATTTCTGTTTGGGAGGAGGAGGAGAAATTTTAACTTTAATAGATATATATA	907
Oy	846	GTCGAGAGAGTTAGAACTCTACAAAGCTCACTTTGGTCCATTCACGTGTGTAGATTTCA	905
Db	908	GTCGAGAGAGATTTAGAACTCTACAAAGGACCTTTTGGTCCATTCACGTGTGTAGATTTA	967
Oy	906	GTCGATGAGGAGAACTCTATAGCAGGAGTCTGGAATGAGGACATATGAGATTTGGCCAA	965
Db	968	GTCGATGAGGAGAACTCTATGCCAGTGTGTTCCAGAAATGGAACATATGAGATTTGGCAA	1027
Oy	966	CTGTGTAGGAAAGACCTATGCGCTGTGGAATTCGTGTNTTCTGTAGAGACACGCGGG	1025
Db	1028	CTGTGTAGGAAAGAACTATGAGCTTTGGAAATGTGTCTTCGTGAAMAAATATGTGTG	1087
Oy	1026	AACGCGAAAGCCAAAGATGCGATTTCCAGAAACAGCAGAGAAAGGTGGCAAAAGAA	1085
Db	1088	AGCTGGCAAAAGCCAAAGATTTGTTTCCAGAGACACAGAGAGAGAGCT---ACAAAGAA	1144
Oy	1086	TTGCTTCAGAAATTCAGATTCATCTATTCATCACTCTGAATTAAGGCTGAGCAT	1145
Db	1145	TTGCTTCAGAAATTCAGATTTGCATTTCTTCCTCACTCTGATGTTAAGGCTCGAGAGCT	1204
Oy	1146	CAGACGTGTGCTGCCGAACCAT--ATGTTCAATGAGCTAAACACAGCAAGCAAGCATCC	1203
Db	1205	CAGATCATTTGTTGCAAGTTAGTATACAATGACTATAAACAGCAACAGAGAAAGCATCA	1264
Oy	1204	GCGCTT--CAGAGTTTACTGTCTGCGCTCGAGCAAGAGGCGAGAAATATTTGGGCAATATGAG	1262
Db	1265	GCGCTTCCAGAGTTACTGTCTGCTTTAAGGCGAGAAACAGCAGTAATATATGAGGAAATATGAA	1324
Oy	1263	TTAGCTCCAGTGCAGCAAGCAAGCT---ACTCAGTGTCCCGGTGAGTGAATTAAGTCTGGA	1317
Db	1325	TTAGCTCCAGTGCAGTGAACAATCACTCACTCACTTGTGTACCTGTAAAGTAAAA---CTC	1380
Oy	1318	GTCGTCGAGAGTGAAGGAGGAGGAGTTATCTCTTTATATAGCAGTGGCTGTATCTTTT	1440
Db	1381	AAGTGTCAAGTGAAGGAGGAGGAGTTATCTCTTTATATAGCAGTGGCTGTATCTTTT	1440
Oy	1374	-----GGATGAAAGGCAACTTACAATCTCAAT--TTACACTTAATATTTCTTTTACT	1426
Db	1441	TATGATATATATACAGGCAACATCCAAATTTCTATTTATTAACAATTTAGGCTGTCTGTACT	1500
Oy	1427	GTTTATGT---TATGAGAGAGAAATATATTTGGCTATTTTCTGATCTTCCCTTAAG	1483

Db	1501	GTATTATGTAAATATGAGAGAGAAACTATATATTTGGCTGATTTTTCGTAT	-----CTTAAG	1556
Qy	1484	AAGATGCCCTTTTGTGCTTCCTCCT--AGTGATCAAGAGAGGAAATACATGATAAAGTAA		1541
Db	1557	CAGATCCCTTTTCTTTCTTTTTCCTTACGTTGTAAAGAACAGGGAATACATGATAAAGTAA		1516
Qy	1542	CCGCTTGATCTCTTCATCTGTACAAAGACGCTGCTTCAGAACGCTCA-TAATTTAGTTA		1600
Db	1617	CTGGTTGATTTTC--TCGTCATCTGTACACCTGCGCTGAACATCTAAATGTTTTAGTTG		1674
Qy	1601	TCTAAATAAATGCTCTAAATTAATAAA		1630
Db	1675	TCTAAATAAATGCTCTAAACAAAAAA		1704

RESULT 5	AB024327	1864 bp	mRNA	linear	PR1 20-DEC-2000
LOCUS	AB024327				
DEFINITION	homo sapiens pt-wd mRNA for WD-40 repeat protein, complete cds.				
ACCESSION	AB024327.1	GI:4519416			
VERSION	AB024327.1	GI:4519416			
KEYWORDS	pt-wd; WD-40 repeat protein.				
SOURCE	homo sapiens liver cell line:hepg2	cDNA to mRNA.			
ORGANISM	homo sapiens				

REFERENCE	AUTHORS	TITLE
1 (sites)	Matsuda, S., Katsumata, R., Okuda, T., Yamamoto, T., Miyazaki, K., Seng, T., Machida, K., Thant, A. A., Nakatsugawa, S. and Hamaguchi, M.	Molecular cloning and characterization of human MMD, a novel MMD-related protein with 40 repeats frequently overexpressed in breast cancer

JOURNAL Cancer Res. 60 (1), 13-17 (2000)
MEDLINE 20110763
REFERENCE 2 (bases 1 to 1864)
AUTHORS Matsuda, S. and Okuda, T.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1999) Satoru Matsuda, Nagoya University School of Medicine, Molecular Pathogenesis; 65 Tsurumai-cho Showa-ku, Nagoya 466-8550, Japan (E-mail: smatsuda@sunu.med.nagoya-u.ac.jp, Tel: 81-52-744-2463, Fax: 81-52-744-2464)
KEYWORDS TCCat100/Qualifiers

Source

gene
CDS

BASE COUN
ORIGIN

Query Match	62.5%	Score 1018.4;	DB 9;	Length 1864;
Best Local Similarity	83.1%;	Pred. No. 5.2e-262;		
Matches 1322;	Conservative	0;	Mismatches 230;	Indels 38; Gaps 13.

QY	261	AAGTGTCTGTTTGGGCGTCACACATTTGATTAAGAGATGCCACAAAGCTGCGACAGAGCTG	320
Db	434	AAGGTGCTGTTTGGGGTGCACACACTGAATTAAGATGCCACAAAGCTACAGACAGCTG	493
QY	321	CAGATTTACAGAGCCAAATATATGGAGTGGGCTCAGAGATGAATTTGATGACCTGGCTC	380
Db	494	CAGATTTACAGAGCCAAATGTGGAGTGGCTCAGAGATGAATTTGATGACCTGGCTC	553
QY	381	ATAGCACATTTGTCAAGACTGTGGATTTTCACACAGGATATAGCAATTAACCTGNTAACTGGGG	440
Db	554	ATMAACACATTTGTCAAGACTGTGGATTTTCACAGAGATATATTAATTTGTTAACCGGGG	613
QY	441	GACAGATTAACCTGTGGCGCATATATGACTTTGAACAAACCTTGAAACAGAACCTTAAGAAA	500
Db	614	GACAGATTAACCTGTAGCCATATATATGACTTTGAACAAACCTTGAAACAGAACCTTAAGAAA	673
QY	501	TCAGTGGCCACACTTGTGATTTAAAAAGCTCTGTGTGGCAGTGACGATTAACAGATTC	560
Db	674	TTAGTGTCTACTTCTGTGATTAAAAAAGCTCTGTGTGGCAGTGAGATTAACAGATTC	733
QY	561	TTTTCAGCGCATATATAAATGTGTGGCTGGGATCATCCCAATATGACAGAAAGTAAAT	620
Db	734	TTTTCGCTATATACAAAACTGTGTGACTTTGGGATCATCTTACTATGACAGAAAGTAAAT	793
QY	621	CTCTGAATTTTAAATATGTCTGTATGACAGCATGGAGTATATTCCTGAAGGAGATTTTGG	680
Db	794	CTCTGAATTTTAAATATGTCTGTATGACAGCATGGAGTATATTCCTGAAGGAGATTTTGG	853
QY	681	TTATTACTTATGAGCATATATTTGCTTTTCATATAGTCACATTAAGTCGTGGAGCCAAATTAAT	740
Db	854	TTATTACTTATGAGCATATATTTGCTTTTCATATAGTCACATTAAGTCGTGGAGCCAAATTAAT	913
QY	741	CCTTTAACACTCCTGAGACCATCAATTCGCGCTCTTTTCATCCAGAGAAAGACTTTCTTG	800
Db	914	CCTTTAAGTCTCTGACACCATCAATTTGCTCATCTCTCACTCGAGAAAAGTTTCTTG	973
QY	801	TTGCGGGTGGAGAAAGCTTTAAACTGTACAAATATATATTAACATGAGGAGAGATTTAG	860
Db	974	TTGCGAGCGGTGAAGATTTTAAACTTTATATAGTATATTAATATGAGGAGAGATTTAG	1033
QY	861	AATCTTACAAAGGCTCACTTGTGTCCATTCACATGCTGTGAGATTCAGTCTGATGGGAAAC	920
Db	1034	AATCTTACAAAGGACACTTTGGTCCATTCACATGCTGTGAGATTTAGTCTGATGGGAAAC	1093
QY	921	TCATAGCCAGCGTTCTGAAAGATGGGACATTTGAGATTTGTGGCAAACTGTGGTGAAGAA	980
Db	1094	TCATAGCCAGTGTCTCGAAAGATGGGACATTTGAGACTATGTGGCAAACTGTGTGAAGAAA	1153
QY	981	CGTATGGCCGTGTGAAATATGCGTGTTCCTCTAGGAGAACACGCGGGAACTGTGGCAAGCCAA	1040
Db	1154	CGTATGGCCCTTTGAAATATGTGTCTTCTCTAGAGAAATATGTGTGAGACTGTGGCAAGCCAA	1213
QY	1041	AGATCGGATTTCCAGAAAACAGCAGAGAGAGAGCTGCCAGAAAGAAATTTGCTTCAGAGATT	1100
Db	1214	AGATTTGTTTTCCAGAGAACAGAGAGAGAGAGCT---AGAAAGAAATTTGCTTCAGAGATT	1270
QY	1101	CAGATTTCCATCTATCAATCAACCTCTGAGTAAAGTGAAGGCTGAGCTCAGAGTGTGCTGC	1160
Db	1271	CAGATTTCCATCTTCTTCACCTCTCTATGTTTAAAGGCTTAGAGCTCAATCATATGTTTGA	1330
QY	1161	GAACGCAT--ATGTTCAATGCACTTAACAGACAGACAGACAGCATCGGCTT--CAGAGTTAC	1217
Db	1331	GTTAGTATTAACAACAGCATTAATAAACAGAGACAGAGAAAAGGATCAGCGCTTCCAAAGTTAC	1390
QY	1218	TGTTGCTGCTGAGGCAAAAGAGGGGCGAAAAATTTGGGGCATTTGAGTTAAGTCCAGTGCAC	1277
Db	1391	TGTTGCTGCTTAAGGCAAAAGCAGTAAATATATAGGAAATATGAATTAAGCTCCAGTGTG	1450
QY	1278	GAAAGCT-----ACTAGTGTTCGCCGTGAGAAATGGCTAGTCAAGTCTCAGAGTGCAG	1332
Db	1451	GAAACACTTAACATACTGTGTGTACGTGTATAGTGAATA-----CTCAAGTGTCAAGTGAAG	1506
QY	1333	GCAGAGAGA-TTGTGCTACATATGATGCCATAGCTGCTGTTT-----GGAATGAAA	1382

[illegible][illegible]

Db 88 CGGACCCGCTGATTTGGCCCTTCAGTGCACATCACGCCCTTAATGGTATTTCTTAATCA 147
 QY 186 GCCTTGCAGAAATGCGAAGCCCATGCTCCGACAGGAGATACAGAGATGGAATGGAA 245
 Db 148 GCGCTGCAGAAATGATGTAACCTATGCTAGCCAGGAGATACAGAGATGGAATGGAA 207
 QY 246 CATTTTGGCTCATTAAGAGTCTCTTTGGGGTGCACATTTGAATTAAGATCCCAAG 305
 Db 208 CATTTTGGGCTCATTAAGAGTCTCTTTGGGGTGCACATTTGAATTAAGATCCCAAG 267
 QY 306 CTGCGACAGAGCTGCAGACTTCACAGCCAAAGTATGGGATGGGCTCCAGAGATGAAT 365
 Db 268 CAGCTACAGAGAGCTGCAGATTTTCACAGCCAAAGTATGGGATGGGCTCCAGAGATGAAT 327
 QY 366 TGAATGACCTTGGCTCATTAAGAGTCTCTTTGGGGTGCACATTTGAATTAAGATCCCAAG 425
 Db 328 TGAATGACCTTGGCTCATTAAGAGTCTCTTTGGGGTGCACATTTGAATTAAGATCCCAAG 387
 QY 426 ACCTGATTAAGAGCTGCAGACTTCACAGCCAAAGTATGGGATGGGCTCCAGAGATGAAT 485
 Db 388 ATTTTGAATTAAGAGCTGCAGATTTTCACAGCCAAAGTATGGGATGGGCTCCAGAGATGAAT 447
 QY 486 CAGAACTTAAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 545
 Db 448 CAGAACTTAAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 507
 QY 546 AGATTAAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 605
 Db 508 AGATTAAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 567
 QY 606 TGACAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 665
 Db 568 TGACAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 627
 QY 666 AAGGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 725
 Db 628 AAGGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 687
 QY 726 TGAAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 785
 Db 688 TGAAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 747
 QY 786 AAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 845
 Db 748 AAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 807
 QY 846 GTGAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 905
 Db 808 GTGAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 867
 QY 906 GTGAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 965
 Db 868 GTGAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 927
 QY 966 GTGAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1025
 Db 928 GTGAGAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 987
 QY 1026 AACTGGCAAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1085
 Db 988 AACTGGCAAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1044
 QY 1086 TTGCTTGAAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1145
 Db 1045 TTGCTTGAAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1104
 QY 1146 CAGAGCTGTCGCTGCGCAAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1204
 Db 1105 CAGAGCTGTCGCTGCGCAAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1164
 QY 1205 CAGAGCTGTCGCTGCGCAAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1263
 Db 1165 CAGAGCTGTCGCTGCGCAAGAGTGCAGACACTTCTGTATTAAGAGCTCTGTGGTGCAGTG 1224

QY 1264 TAGCTCAGTGCAGAGAGCT-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1318
 Db 1225 TAGCTCAGTGCAGAGAGCT-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1284
 QY 1319 TGTCTGAGTGCAGAGAGAGCT-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1375
 Db 1285 TC-----AGATGAAGAGAGAGTGCAGAGCT-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1340
 QY 1376 -----AATGAAGAGAGAGTGCAGAGCT-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1427
 Db 1341 AATGAAGAGAGAGTGCAGAGCT-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1400
 QY 1428 TTTATGCT-----TATGAAGAGAGAGTGCAGAGCT-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1484
 Db 1401 TTTATGCT-----TATGAAGAGAGAGTGCAGAGCT-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1456
 QY 1485 AGATGCTGTTGCCCTGAGTGAAGATGGCTAG-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1542
 Db 1457 AGATGCTGTTGCCCTGAGTGAAGATGGCTAG-----ACTCAGCTGTTGCCCTGAGTGAAGATGGCTAG 1516
 QY 1543 CGGTTGA--TCTCTTCAATGTTACAGAGCT 1572
 Db 1517 TGGTTGATTTCTCGTTCAATGTTACAGAGCT 1548

RESULT 8
 AC010878/c
 LOCUS AC010878 154115 bp DNA linear PRI 07-NOV-2001
 DEFINITION Homo sapiens BAC clone RP11-230E20 from 2, complete sequence.
 ACCESSION AC010878
 VERSION AC010878.4 GI:13346585
 KEYWORDS HTG.
 SOURCE
 ORGANISM Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 154115)
 Sulston, J.E. and Waterston, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 9847074
 PUBMED
 REFERENCE
 2 (bases 1 to 154115)
 Sandholzer, T., Abbott, A., Doeber, A. and Markovic, C.
 The sequence of Homo sapiens BAC clone RP11-230E20
 Unpublished
 3 (bases 1 to 154115)
 Waterston, R.H.
 Direct Submission
 Submitted (25-SEP-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 154115)
 Waterston, R.H.
 Direct Submission
 Submitted (15-MAR-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 154115)
 Waterston, R.
 Direct Submission
 Submitted (09-AUG-2001) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 6 (bases 1 to 154115)
 Waterston, R.
 Direct Submission
 Submitted (07-NOV-2001) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 15, 2001 this sequence version replaced gi:11527493.

COMMENT
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics -----
 Center project name: H_NH0230E20

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is RP11-278G12, 200 bp overlap; the clone sequenced to the left is RP11-92F3, 200 bp overlap. Actual start of this clone is at base position 1 of RP11-230E20; actual end is at base position 23434 of RP11-278G12.

RP11-230E20 contains polymorphisms with RP11-278G12.

Location/Qualifiers

1. 154115

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone_lib="RPCT-11"

1. .97

/rpt_family="L1"

158. .389

/rpt_family="MIR"

450. .788

/rpt_family="L1"

1032. .1326

/rpt_family="Alu"

1744. .1781

/rpt_family="MER1_type"

1735. .2026

/rpt_family="Alu"

2032. .2132

/rpt_family="MER1_type"

2164. .2252

/rpt_family="L1"

2258. .2613

/rpt_family="L1"

2633. .2895

/rpt_family="Alu"

2899. .2919

/rpt_family="TAA)n"

2923. .3480

repeat_region

/rpt_family="L1"

3552. .3659

/rpt_family="L1"

3670. .3777

/rpt_family="L1"

4243. .4534

/rpt_family="Alu"

4827. .5132

/rpt_family="Alu"

5718. .5925

/rpt_family="MER1_type"

6210. .6792

/rpt_family="L2"

6933. .6994

/rpt_family="ERV1"

6995. .7535

/rpt_family="ERV1"

7516. .7619

/rpt_family="ERV1"

7620. .7921

/rpt_family="Alu"

7922. .8073

/rpt_family="ERV1"

8078. .8386

/rpt_family="Alu"

8456. .8893

/rpt_family="ERV1"

9675. .9781

/rpt_family="MALR"

10536. .10774

/rpt_family="MALR"

10784. .11233

/rpt_family="ERV1"

11428. .11544

/rpt_family="MALR"

11545. .11899

/rpt_family="MALR"

11900. .11920

/rpt_family="MALR"

12078. .12359

/rpt_family="Alu"

12420. .12706

/rpt_family="MALR"

14365. .14403

/rpt_family="AT-rich"

14399. .14528

/rpt_family="MER1_type"

14930. .15505

/rpt_family="ERV1"

15849. .16157

/rpt_family="Alu"

17559. .17692

/rpt_family="MIR"

18143. .18624

/rpt_family="ERV1"

18625. .19034

/rpt_family="ERV1"

19035. .19080

/rpt_family="ERV1"

19614. .19860

/rpt_family="MIR"

20803. .20906

/rpt_family="MER1_type"

20994. .21169

/rpt_family="L1"

21214. .21402

/rpt_family="L1"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

QY	959	TGGCAAACTGTGGTAGGAAAGACCTATGCGCTGTGGAAATGCGTGNTTCTGTAGGAAGAC	1018
Db	94015	GGCGAAACTGTGTGTAGGAAACACATACGCGCTTTGGAAATGTGTCTGTGAAGAAAG	93956
QY	1019	AGCGGGCAACTGGCAAAAGCCAAAGATCGAGATTTCTCAGAAACAGACAGAGAAAGACTGTGCA	1078
Db	93955	AGTGTGTAGCTGTGCAAAAGCCAAAGATGTTGGCTTCCAGACACACAGAAAGAGACT---	93899
QY	1079	GAATAAATTTGTTTACAGAAATTCAGATTCCATCTATTTCATCACTCTGTGAAGTTAAGCC	1138
Db	93898	GAATAAATTTGTTTACAGAAATTCAGATTCCATCTATTTCATCACTCTGTGAAGTTAAGCC	93839
QY	1139	TGAGCATCAGACGCTGTCTCTCCCAAAACCAT-ATGTGTAGACTAA- ----CAAGAGA	1192
Db	93838	TGAGGTCATATCATATGCTCTGAGTTACTATTCACAACTATGTGACTTAAACAGACAGAGAGA	93779
QY	1193	GACACAGATCCGCTT-CAAGATTTACTGTCTGCTGAGGCAAAAGAGGCGAAATATTG	1251
Db	93778	GAATAAGCATAGCCCTTCCAAAGTACTCTGCTTAAAGCAGAAAGAGCAGTAAATATG	93719
QY	1252	GGGCAATATGATTAGTCCAGTCCAGCAACAGCT-ACTCAGTTGTGCGCTGAGTAAAA	1310
Db	93718	GGGCAATATGATTAGTCCAGTCCAGTCCAGCAACAGCT-ACTCAGTTGTGCGCTGAGTAAAA	93659
QY	1311	TGCGTGAGTGTCTGAGGTGCGAGCAGAGAGATTTGTGCTCATATAGTCCATACGCTGCTG	1370
Db	93658	CTCAAGAGTATCAGATGAAGAGAGAG-GTAGGTATATCTCTGTGTAAAGTACAGTGGCCGTGTA	93600
QY	1371	TTTGG-----AATGAAAGACCAATTCACATCTCCATT-TTACACCTAAATTTCTT	1420
Db	93599	TCTTTTAAATGAATATATACAGCCCAACATCCATTTCTTAAATACATTAAGGGTCTT	93540
QY	1421	TTAGCTGTTTATG--TTATGAAAGAAATAATATTTGGCTATTTTCTGAC--TTTC	1475
Db	93539	GTACGCTTTTATTTATTTATTTATGAGAGAAATAATATTTGGCTATTTTCTGACTTT	93480
QY	1476	CCTTAAGAAGATGCTTTTGTCTCTGTC-CTAGTGAAGAGAGAGAAATACATGAT	1534
Db	93479	CCTTAAGAAGATGCTTTTGTCTCTGTC-CTAGTGAAGAGAGAGAAATACATGAT	93420
QY	1535	AAAGTAACGGTTTGATCTCTTTCATTTGACAAAGCAGCTCTCAGAACAGCTCA-TATTT	1593
Db	93419	AAAGTAACGGTTTGATCTCTTTCATTTGACAAAGCAGCTCTCAGAACAGCTCA-TATTT	93362
QY	1594	TTAGTATCTAATAATAAAGCCTCTAATAATAAAAAA	1630
Db	93361	TTAGTATCTAATAATAAAGCCTCTAATAATAAAAAA	93325
RESULT 9			
AC010132/c			
LOCUS	AC010132	166680 bp	DNA
DEFINITION	Homo sapiens BAC clone Rp11-11K18 from 7p11.2-p2, complete	linear	PRI 30-SEP-2000
ACCESSION	AC010132		
VERSION	AC010132.5	GI:10440742	
KEYWORDS	HTG.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteleostomi; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 166680)		
JOURNAL	Sulston,J.E. and Waterston,R.		
MEDLINE	Toward a complete human genome sequence		
PUBMED	Genome Res. 8 (11), 1097-1108 (1998)		
REFERENCE	9847074		
AUTHORS	2 (bases 1 to 166680)		
TITLE	Andrews,S., Wohlmann,P. and Scherger,E.		
JOURNAL	The sequence of Homo sapiens BAC clone Rp11-11K18		
REFERENCE	3 (bases 1 to 166680)		

AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 166680)
Waterston, R.

AUTHORS Direct Submission
TITLE Submitted (30-SEP-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
JOURNAL On Sep 30, 2000 this sequence version replaced gi:8954217.

COMMENT

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@wustl.edu

----- Summary Statistics

Center project name: H_NH011K18

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:mallo:green@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatem, M., Catanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is GS1-308H5, 200 bp overlap. Actual start of this clone is at base position 156485 of GS1-308H5; actual end is at base position 166680 of RP11-11K18.

There are polymorphic base pair differences in the overlap between the clone RP11-11K18 and GS1-308H5.

FEATURES

SOURCE

Location/Qualifiers

1..166680

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="7"

/map="7p11.2-p2"

/clone="RP11-11K18"

/clone_11b="RPCT-11"

2..397

/rpt_family="ERV1"

repeat_region

1478..1785

/rpt_family="Alu"

repeat_region

repeat_region
2138..2508
/rpt_family="ERV1"
repeat_region
2532..2817
/rpt_family="MALR"
repeat_region
2846..2911
/rpt_family="ERV1"
repeat_region
2912..3043
/rpt_family="Alu"
repeat_region
3044..3115
/rpt_family="ERV1"
repeat_region
3087..3596
/rpt_family="ERV1"
repeat_region
3600..3780
/rpt_family="ERV1"
repeat_region
3781..4090
/rpt_family="Alu"
repeat_region
4091..4257
/rpt_family="ERV1"
repeat_region
4260..4865
/rpt_family="ERV1"
repeat_region
4902..4994
/rpt_family="ERV1"
repeat_region
6077..6610
/rpt_family="ERV1"
repeat_region
7185..7327
/rpt_family="ERV1"
repeat_region
7327..7539
/rpt_family="ERV1"
repeat_region
7579..8074
/rpt_family="ERV1"
repeat_region
8169..9081
/rpt_family="ERV1"
repeat_region
9082..9362
/rpt_family="Alu"
repeat_region
9363..9633
/rpt_family="ERV1"
repeat_region
9634..9943
/rpt_family="Alu"
repeat_region
9944..10150
/rpt_family="ERV1"
repeat_region
10151..10567
/rpt_family="MALR"
repeat_region
10568..12227
/rpt_family="ERV1"
repeat_region
12316..12514
/rpt_family="ERV1"
repeat_region
12532..12582
/rpt_family="ERV1"
repeat_region
13068..13614
/rpt_family="ERV1"
repeat_region
15167..15445
/rpt_family="Alu"
repeat_region
16828..17212
/rpt_family="MALR"
repeat_region
17548..17857
/rpt_family="Alu"
repeat_region
17897..18030
/rpt_family="ERV1"
repeat_region
18365..18652
/rpt_family="Alu"
repeat_region
19118..19746
/rpt_family="ERV1"
repeat_region
19750..19814
/rpt_family="ERV1"
repeat_region
19852..20169
/rpt_family="ERV1"
repeat_region
20224..20412
/rpt_family="ERV1"
repeat_region
20540..20840
/rpt_family="Alu"
repeat_region
20879..20997
/rpt_family="MIR"


```

repeat_region      21649..21848
                    /rpl_family="MERL_type"
repeat_region      22158..22177
                    /rpl_family="MIR"
repeat_region      22178..22488
                    /rpl_family="ERV1"
repeat_region      22489..22777
                    /rpl_family="Alu"
repeat_region      22778..22919
                    /rpl_family="ERV1"
repeat_region      22920..22979
                    /rpl_family="MIR"
repeat_region      23031..23067
                    /rpl_family="L2"
repeat_region      23518..23928
                    /rpl_family="MALR"
repeat_region      24091..25047
                    /rpl_family="L2"
repeat_region      25094..25312
                    /rpl_family="L2"
repeat_region      25375..25349
                    /rpl_family="L2"
repeat_region      25663..25817
                    /rpl_family="MERL_type"
repeat_region      27098..27212
                    /rpl_family="L2"
repeat_region      27657..27925
                    /rpl_family="MALR"
repeat_region      28020..28061
                    /rpl_family="L2"
repeat_region      28527..28780
                    /rpl_family="Alu"
repeat_region      29058..29361

```

Query Match 49.9% Score 812.8; DB 9; Length 166680;
 Best Local Similarity 78.9%; Pred. No. 1.6e-206;
 Matches 1179; Conservative 0; Mismatches 270; Indels 46; Gaps 16;

```

QY 163 GCCTTACGCTACTTTCATGATCAGCGCTTGCAGAGATGCGACCCATGCTCCGCGAGG 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48564 GCTTGTAGATTTCTTAAATCAGCTTGCAGAGATGCGACCCATGCTCCGCGAGG 48505

QY 223 AGATACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48504 AGATACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 48445

QY 283 ATGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48444 ACTGAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 48389

QY 343 GGATGCGGTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 402
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48388 GGATGCGGTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 48329

QY 403 GGATGCGGTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48328 GGATGCGGTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 48269

QY 463 ATATGACTTGAACAACTGAGAGAACTGAGAGAACTGAGAGAACTGAGAGAACTGAGAGAA 522
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48268 ATATGACTTGAACAACTGAGAGAACTGAGAGAACTGAGAGAACTGAGAGAACTGAGAGAA 48209

QY 523 TAAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48208 TAAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 48149

QY 583 TCGG-CTCTGGGATCATGCCAATGACAGAGAACTGAGAGAACTGAGAGAACTGAGAGAA 641
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48148 TAAAGACTTGGGATCATGCCAATGACAGAGAACTGAGAGAACTGAGAGAACTGAGAGAA 48089

QY 642 TTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 701
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48088 CTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 48029

```

```

QY 702 TTGCTTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48028 GTGCTTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 47969

QY 762 TCAATTCGCTCTTTCATCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 821
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47968 TTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 47909

QY 822 AACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47908 AACT-TATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 47850

QY 882 GTCCATTCACCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 941
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47849 GTCCATTCACCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 47790

QY 942 ATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47789 ATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 47732

QY 1002 TGNCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1061
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47731 TGNCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47672

QY 1062 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47671 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47618

QY 1122 CTCTGTAAGTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47617 CTCTGTAAGTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47558

QY 1179 ACTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47557 ACTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47498

QY 1238 GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47497 GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47438

QY 1297 CCCGTGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47437 CCCGTGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47379

QY 1357 GCCATAGCCTG--CTGTTGGAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47378 GCCATAGCCTG--CTGTTGGAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47319

QY 1406 ACACTTAATTTCTTTAGCTGTTATG--TTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1462
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47318 ACACTTAATTTCTTTAGCTGTTATG--TTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47259

QY 1463 TTTTCTGAC--TTTCCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1515
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47258 TTTTCTGAC--TTTCCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47199

QY 1516 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1575
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47198 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47143

QY 1576 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1629
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47142 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 47088

RESULT 10
AC010101/c AC010101 187017 bp DNA linear HTG 08-AUG-2000
LOCUS Homo sapiens chromosome 7 clone RP11-447C21, WORKING DRAFT
DEFINITION
SEQUENCE, 8 unordered pieces.
ACCESSION AC010101
VERSION AC010101.8 GI:9739351

```

KEYWORDS
SOURCE
ORGANISM

HTG: HTGS_PHASE1; HTGS_DRAFT.

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

(bases 1 to 187017)

Waterston, R.H.

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 187017)

Waterston, R.H.

Direct Submission

Submitted (11-SEP-1999) Genome Sequencing Center, Washington

University School of Medicine, 444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Aug 8, 2000 this sequence version replaced gi:9653201.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: MUGSCWeb site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH047C21

----- Summary Statistics -----

Sequencing vector: M13; 63%

Sequencing vector: plasmid; 37%

Chemistry: Dye-terminator Big Dye; 52% of reads

Chemistry: Dye-terminator Big Dye; 52% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 18198 bases at least Q40

Consensus quality: 183284 bases at least Q30

Consensus quality: 184219 bases at least Q20

Insert size: 190000; agarose-fp

Insert size: 186317; sum-of-ctrls

Quality coverage: 6.29 in Q20 bases; agarose-fp

Quality coverage: 6.62 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.* 1 4028: contig of 4028 bp in length
* 4029 4128: gap of unknown length
* 4129 7967: contig of 3839 bp in length
* 7968 8067: gap of unknown length
* 8068 16351: contig of 8284 bp in length
* 16352 16451: gap of unknown length
* 16452 26437: contig of 9986 bp in length
* 26438 26537: gap of unknown length
* 26538 35888: contig of 9351 bp in length
* 35889 35988: gap of unknown length
* 35989 56349: contig of 20361 bp in length
* 56350 56449: gap of unknown length
* 56450 85530: contig of 29081 bp in length
* 85531 85631: gap of unknown length
* 85631 187017: contig of 101387 bp in length.

FEATURES

source

1. 187017

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="7"

/clone="Rp11-447C21"

1. 4028

/note="assembly_name:Contig26"

4129. 7967

/note="assembly_name:Contig27"

8068. 16351

/note="assembly_name:Contig28"

16452. 26437

misc_feature

misc_feature

misc_feature

/note="assembly_name:Contig29

clone_end:17

vector_side:left"

26538. 35888

/note="assembly_name:Contig30"

35989. 56349

/note="assembly_name:Contig31"

56450. 85530

/note="assembly_name:Contig32

clone_end:SP6

vector_side:right"

85631. 187017

/note="assembly_name:Contig33"

BASE COUNT 53793 a 37172 c 37140 g 58180 t 732 others

ORIGIN

Query Match 49.9% Score 812.8; DB 2; Length 187017;

Best Local Similarity 78.9% Pred. No. 1.6e-206;

Matches 1179; Conservative 0; Mismatches 270; Indels 46; Gaps 16;

QY 163 GCCTTACGCGTACTTCTGATCAGCGCTTGCACAAAGATGGCAGCCGATGCTCCGCGAGG 222

Db 100203 GCTTTATGAGATTTCTTAATTCAGTCTGCACAAAGCGCAACCTATGCTATGCCAGG 100144

QY 223 AGATACAGGAGACTGATTTGGAACATTTTGGGTCAATAAGGTCGTGGGTGCAC 282

Db 100143 AGATACAGGAGACTGATTTGGAACATTTTGGGTCAATAAGGTCGTGGGTGCAC 100084

QY 283 ATGATACAGGAGACTGATTTGGAACATTTTGGGTCAATAAGGTCGTGGGTGCAC 342

Db 100083 ACTAATACAGGAGACTGATTTGGAACATTTTGGGTCAATAAGGTCGTGGGTGCAC 100028

QY 343 GGATGCGGTCTCAGAGAGATTTGATGACCTGCTCATTAACGATTTGCAAGACTGT 402

Db 100027 GGATGCGGTCTCAGAGAGATTTGATGACCTGCTCATTAACGATTTGCAAGACTGT 99968

QY 403 GGATTTACACAGATATGCAATTTACCTGATTTGAGGAGAGATTTGAGGAGATTT 462

Db 99967 GGATTTACACAGATATGCAATTTACCTGATTTGAGGAGAGATTTGAGGAGATTT 99908

QY 463 ATATGACTTGAACCAACCTGAGAGAGAGATTTGAGGAGAGATTTGAGGAGATTT 522

Db 99907 ATATGACTTGAACCAACCTGAGAGAGAGATTTGAGGAGAGATTTGAGGAGATTT 99848

QY 523 TAAAGGCTCTGCTGCTGCAAGATTAACAGATTTTCAAGGAGATTTTCAAACTGT 562

Db 99847 TAAAGGCTCTGCTGCTGCAAGATTAACAGATTTTCAAGGAGATTTTCAAACTGT 99788

QY 583 TCGG-CCTGCGATCATGCAACATTAACAGAGATTTTCAAGGAGATTTTCAAACTGT 641

Db 99787 TCGG-CCTGCGATCATGCAACATTAACAGAGATTTTCAAGGAGATTTTCAAACTGT 99728

QY 642 TTACAGAGATTAATTTCTGAGAGAGATTTTCTGATTTTCTGATTTTCTGATTTT 701

Db 99727 TTACAGAGATTAATTTCTGAGAGAGATTTTCTGATTTTCTGATTTTCTGATTTT 99668

QY 702 TTGCTTTTCAATGAGCAATTTCTGAGAGATTTTCTGATTTTCTGATTTTCTGATTTT 761

Db 99667 TTGCTTTTCAATGAGCAATTTCTGAGAGATTTTCTGATTTTCTGATTTTCTGATTTT 99608

QY 762 TCAATTCGCTGCTGCTGCTGCAAGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTT 821

Db 99607 TCAATTCGCTGCTGCTGCTGCAAGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTT 99548

QY 822 AACTGATCAAGATTTTATTAACAGAGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTT 881

Db 99547 AACTGATCAAGATTTTATTAACAGAGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTT 99489

QY 882 GTCCATTAAGATTTTATTAACAGAGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTT 941

Db 99488 GTCCATTAAGATTTTATTAACAGAGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTT 99429

QY 942 ATGGACATTAAGATTTTATTAACAGAGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTT 1001

Accession	LOCUS	DEFINITION	Accession	LOCUS	DEFINITION
AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.	AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
AL645802.11	GI:19572035	HTG.	AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
house mouse.			AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
Mus musculus			AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
Leongamornliert.D.			AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
Direct Submission			AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail: enquiries@hinxton.sanger.ac.uk			AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
On Mar 21, 2002 this sequence version replaced gi:19335832.			AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.			AL645802	AL645802/c	Mouse DNA sequence from clone Rp23-12809 on chromosome 11, complete sequence.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-128D9 is from the RPO-1/3 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6.

BASE COUNT	64093 a	48521 c	50155 g	63751 t
ORIGIN				

Query Match	41.4%	Score 674.2	DB 10	Length 226520
Best Local Similarity	81.2%	Pred. No. 2.0e-169		
Matches 951; Conservative	0	Mismatches 152	Indels 68	Gaps 12

61 CCTCTCTCCCTCCCTCCCTGTCGGCCATGGCCATGAGCCAGACGCCGCTCACTTGTCTGGG 120

Db 142468 CATTGACCTCAGTCCCTGCTCCGCGTGGCCATGAGCCAGACTTCTGCTCATTGGTCAAG 142409

121 CCACACGCGCCCCGTGCTGGATTGCGCCCTTCAAGCGCATCAGCGCTTAAGGCTACCTTTTCCT 190

[illegible]

181 GATACGCGCCCTTTCAGCATTGCCTTTGGGTCTTGCAATGAGGCTGGCTTTCT

142349

db_142348 CATTGACCTCCAGGATACAGGACTGGAT 240

142289 GATCAGTGCCTGCAAGATGCGCAACCTATGCTCCACCAGGAGATACAGGAGACTAGAT

241 TGGACACA-TTTTGGGTCATAAAGGTGCTGTTTGGGGTGCAACATTG---AATAAGGATG 296

Db 142288 TGGACATTTTTTGGGTCATTAAGGTCGTTTGGGGTACAACATTTGACTAATAAAGATG 142229

297 CCACCAAGCTGCGACAGCAGCTGCAGACTTCACAGCCAAAGTATGGGATGCCGCTCTCAG 356

Db 142228 CCAACAAACTGCTACAGCAGCTGCAGACATCAC-----AGTGTGTATGCTGTCTAG 142175

357 GAGATGAATTGATGACCCCTGGCTCATAGCAGCATTTGTCAAGACCTCTGCAATTTGACCAACCAAC 416

[illegible]

ATAGCGAATTGACCTCAATTATGTCCGGGAGTAAAGTTTCACACACAGG

1417 142115

[illegible]

DB 142114 AATGACGTTACCTGCTGACT-GGGCACAGGATAAAGTGTTCACATACACGACTTGA-A 142057

477 AACCTGAAGCAGAACCTTAAGGAATCACTGGCCACACTTCTGGTATTAAAAAGGCTCTGT 536

Db_142056 AACCTGAAGCAGAACCTAAGGAATTAATGCTCACACTTCGGGTATTAAAAAGCCTCTGT 141997

537 GGTCAGTGCAGATAACAGATCCCTTTTCAGCGGATGATTAACCTGTTCCGCTCTGGGATC 596

Db 141996 GGTGCAGTGAAGATAATAGTTCCTTCAGCTGATGATGATAAACTGTTTGGCTTGGCATC 141937

597 ATGCCACATGACAGAACTGAATCTGTGAATTATATATGTCCTGGTACCGAACCACCC

Ddb 141936 ATATCCACAGTCAGCACTGCGAAACTC-
||||| ||||| | | ||||| | |

036

657 ATTATTTTCTGACAGCAGCGATTGGGCTGTTTGCCGAGCAGCAAAAT 141879

[illegible]

Up 1418/8 AATTTCCTGAGGAGAGATCTTGGTATTGCTTTTCATGTGAGTATTATTATTGTC 141819

OY	537	GGGACAGTACATATAAACAGATCCCTTTACGGATGTATTAACCTGTTGGCTCGGCATC	596
Db	101381	GGTGCACTAATAAATAAGATCCCTTCACGGGATGTATTAACCTGTTGGCTCGGCATC	101440
OY	597	ATGCCAAATGACACAAGATGAATAATCTCGAATTTTAATATGTCGTAGACATGACT	656
Db	101441	ATACCAAGATGACACAAGATGAAATCT -TMAATATTAATATGCT -TTACGACAGCAAAT	101498
OY	657	ATTATTCCTAAGAGAGATTTTGGTTATTACTTAATGACAGCATTAATGCTTTTCATATG	716
Db	101499	AATATTCCTAAGAGAGATTTTGGTTATTAAC -TTTTCCGACTTTGGTT -TTCCGACTTTGGTT	101544
OY	717	CAGTAGTGTGGAGCAATTAATCCCTTAAGCTCTCGGACCATCAATTTCCGCTCN	776
Db	101545	ATTAAAGCCCCGAGCAATTAATCCCTTAAGCTCTCGGACCATCAATTTCCGATCTC	101604
OY	777	TTCATCCAGAGAAGAGATTTCTGTGCGGGTGGAGAAAGCTTAATCATGTCAGATNG	836
Db	101605	TTTCATGTAGAGAGAAGATTTTTTTTGTTCGACGTGGAGAAAGCTTTAAACGTATTAAGTACT	101664
OY	837	ATTATTAACAGTGGAGAAGAGATTAATCTTACAAAGGTCACTTTGGTCCATTCAGCTG	896
Db	101665	TTTATTAAGTGGAGAAAGATTAATCAATCCATCAAAAGTCAATTTGGTCTCATTTCA -GTG	101722
OY	897	TGAGATTCAGTCTGATGGGGAACCTATCCACGCGGTGTGAAGTGGGACATTTAGAT	956
Db	101723	TGAGATTCAGTCTGATGGGGAACCTATCCATCCAGCTGTTTCTGAAGTGGGACCCGAGAT	101782
OY	957	TGTGGCAACCTGTGTGAGAAAGACCTATNGCCTGTGAAATGCGTGMTTCTTGAGSAG	1016
Db	101783	TGTGGCAACCTGTGTGAGAAAGACCTATNGCCTGTGAAATGCGTGMTTCTTGAGSAG	101842
OY	1017	ACAGCGGGGAACCTGSCAAACGCCAAAGTCGGATTTCCAAAACACAGAGAAGAGCTGG	1076
Db	101843	ACAGTGGGGGAACCTGSCAAACGCCAAAGTTCGATTTCCAAAACACAGAGAAG-----GC	101896
OY	1077	CAGAAGAAATTCCTCACAGAAATTCAGATTCATCTATTCATCACTCTGAAGTTAAG	1136
Db	101897	TGGAAAGAAATTAATTCAGAG-----AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG	101920
OY	1137	CCTGAGCATCAGACGTGTGCTGCGCAACCATAT -GTTCATGAGCATTAACAGCAGAGAC	1195
Db	101921	CCGAGCATCAAAATGTGTGTGCCAAAGATACCATTTAGTTCATGAGCATTAATTAACCGAGAC	101980
OY	1196	AAGCATCCGCTTCAGAGATTACTGTCTGCT 1226	
Db	101981	AAGCATCCGCTTCAGAGATTACTGTCTGCT 102011	
RESULT_14			
AFJ315726			
LOCUS	AFJ315726	1829 bp	mRNA linear VRT 06-FEB-2001
DEFINITION	Carassius auratus gibelio serine-threonine kinase		
ACCESSION	AFJ315726		
VERSION	AFJ315726.1	GI:12667269	
KEYWORDS			
SOURCE			
ORGANISM	Carassius auratus gibelio.		
	Carassius auratus gibelio.		
	Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;		
	Actinopterygii; Neopterygii; Teleostei; Ostariophysi;		
	Cypriniformes; Cyprinidae; Carassius.		
REFERENCE	1 (bases 1 to 1829)		
AUTHORS	Xie,J., Wen,J.J. and Gult,J.F.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-OCT-2000) State Key Laboratory of Freshwater Ecology		
	and Biotechnology, Institute of Hydrobiology, Chinese Academy of		
	Sciences, Lojia Street, Wuhan, Hubei 430072, P.R.China		
FEATURES			
source	1..1829		
	/organism="Carassius auratus gibelio"		
	/db_species="gibelio"		
	/db_xref="taxon:.101364"		

/db_xref="GI:16768658"
 /translation="MAANNLRQIPITCGSHRPVYHLPDSDICDAGYFLISACKDGP
 MLRHIDGDMVTEGEGKAVNATNLNAGNATLNPAPAPREYAGHGAIRALFCKGDK
 KHIYKSVAFDSENIYVGSNEKLYRVNLEOPAPREYAGHGAIRALFCKGDK
 IISAEDKTVLRMRTGIEYORLOFNPNPISLEISDNHLLTISHSSISFWEIDTL
 KRIKLEVPYPTVASASLHPDKHVFVCGEDPKMKFDYITGNETLEISFEKGHGPAHSAVK
 FSPDGEILASSEDEGTELRKMTQTVGKTELMKCTEPADLGNSSISPREVQAN"
 BASE COUNT 436 a 451 c 425 g 279 t
 ORIGIN

Query Match 20.7%: Score 338; DB 3; Length 1591;
 Best Local Similarity 60.3%; Pred. No 2.3e-79;
 Matches 573; Conservative 0; Mismatches 374; Indels 3; Gaps 1;

QY 68 CTTCCCTCCCTGTCGCCATGAGGAGAGCCGCTACTTGTCTGGCCACAG 127
 DB 336 CCCACTGCGATGCGCCGACACACTTGGACAAATCCCGTACCTGACATGGCCACAG 395
 QY 128 CGGCCGCTGTGATNTGGCTTCAGGGCATCAAGCTTACGGCTACTTGTATCAGC 187
 DB 396 CGACCCGCTGTGACACTTCAAGCAGATGCGACGCGGCTACTTCTCATCTCG 455
 QY 188 GCTTGCAAAAGTGCMAAGCCATGCTCCGCGAGAGATACAGAGACTGATGGAACA 247
 DB 456 GCTTGCAAAAGTGCMAAGCCATGCTCCGCGAGAGATACAGAGACTGATGGAACA 515
 QY 248 TTTTGGGTCATAAAGGTCGTTGGGGTGCACATTGAATGAATGCCACCAAGCT 307
 DB 516 TTGAGAGGACAAAGGCGCGGTGTGAAAGCCGCTGAACCAATGCCACCTGGCC 575
 QY 308 GCGACACAGCTGCACACTTACAGCCAAAGTATGGATGCGGTCTCAGAGATGATG 367
 DB 576 GCGTCCGAGCGCGCCACTTACCGGCAAGTGTGGAATGCGGTGACCGGGCCGAGATA 635
 QY 368 ATGACCTGGCTCATTAAGCATTTGCAAGACTGTGATTTCAACAGATAGCAATTAC 427
 DB 636 CACAGCTTCCAGACACAGCATGCAATGCAAGGCTGCTTCGACAGAGATTCGAGAAC 695
 QY 428 CTGNTAAGTGGGAGAGATTAAGTCTGCGATATGACTTGAACCAACCTGAAGCA 487
 DB 696 ATAGTACTGCGACAGCAAGGATGTGCGGCTTTCACACTGAGAGCCCGAGGGCG 755
 QY 488 GAACCTAAGGAATCAGTGGCCACACTTCTGATTAATAAAGCTCTGTGTCAGTAC 547
 DB 756 CAGCCGAGAGATGACCGCGACATACGCGGCGCATCAAGCGGCGCTTCTGCGCGGCG 815
 QY 548 GATTAACAGATCTTTCACAGC---GATGATTAACCTGCTGCGATCATGCGACA 604
 DB 816 GACAAAGTGCATCATCTCGGCGCGAGCAAGAGGTGCGCTTGGGACCGCATGACG 875
 QY 605 ATGACAGAAAGTGAATCTGCAATTTAATATGCTGTAGCAGCATGAGATATTCCT 664
 DB 876 GGCATCGAGTGCAGCGCTGAGTTCAACAGCAATCCGACAGCCTAGAGATCTCCAGC 935
 QY 665 GAAGAGAGATTTGTTTACTTATGAGAGATCTATTGCTTTTCATAGTGCAGTAAGT 724
 DB 936 GACACACATATCATCATATGACAGCGCTCTGATTAAGCTTCTGGAGATCGACAGC 995
 QY 725 CTGGAGCCATTAATTCCTTGAAGCTCTGTGACCATTAATTCGCTTNTTCATCCA 784
 DB 996 CTGAAGAGCTGAAGAGGTGAAGTCCGAGCAAGCGTGGCGCCAGTTTGCATCCG 1055
 QY 785 GAGAGAGATTTCTGTTGGGCTGAGAAAGCTTTAACTGTACAGATGATGATTAAC 844
 DB 1056 GATAAGCATGTGTTGTTGTTGGGCGAGAGACTTAAAGATGTACAAATTCGACTACATC 1115
 QY 845 AGTGAGAGAGTGAAGTCTTACAAAGTCACTTGTGCTCATTCAGTGTGAGATTC 904
 DB 1116 ACAGGAAACGAATGATCTTCAAGGACATTTGGGCGCGTGCACAGCGTGAAGTTC 1175
 QY 905 AGTCTGATGGGAACTCTATGCAAGCGGTTCTGAAGTGGACATTAAGATTTGGCAA 964

DB 1176 AGCCGAGATGGCGAAGTATGCAAGCGGCTCCGAGAGACGACACTTCCGGCTGTGACAG 1235
 QY 965 ACTGTGTAGGAAGAGACCTATGCGCTGTGAAATGCGTGTGTTCTGAGGA 1014
 DB 1236 ACCACCGTGGCGAAGACCTAGCGCTTGTGGAAGTGCACAGGACCGCGGGA 1285

Search completed: February 6, 2003, 14:53:09
 Job time : 6115 secs

1
2
3
4
5

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100